Modelling in Medical Technology Assessment

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Michel, Bowine Christine

Modelling in Medical Technology Assessment / Bowine
Christine Michel. - [S.1. : s.n.] (Ridderkerk : Ridderprint). - Ill.
Thesis Erasmus Universiteit Rotterdam. - With ref. - With summary in Dutch.
ISBN 90 - 9009061 - 4
NUGI 743
Subject headings: Medical Technology Assessment; modelling /
health care; economic aspects.

Modelling in Medical Technology Assessment

Modellering in Medische Technology Assessment

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus Prof dr PWC Akkermans MA
en volgens het besluit van het college voor promoties.
De openbare verdediging zal plaatsvinden op
woensdag 24 januari 1996 om 15.45 uur

door

Bowine Christine Michel

geboren te Leiden

Promotiecommissie

PROMOTOR:	Prof dr FFH Rutten
CO-PROMOTOR:	dr BA van Hout
OVERIGE LEDEN:	Prof dr PJ van der Maas dr HR Büller dr J Kievit

Cover:	Sculpture by Elly Hendrix, Rotterdam-Delfshaven
Photography:	Max Koot Studio, the Hague
Printed by:	Ridderprint, Ridderkerk

"Why," said the Dodo, "the best way to explain it is to do it."

Lewis Carroll: Alice in Wonderland

Contents

INTRODUCTION

Health care is a rapidly developing field in which new technologies are introduced continuously. Not all new technologies have the same impact however: most represent only small changes in existing technologies, whereas only a few - like organ transplants - really are revolutionary new developments. Some of these new technologies eventually lead to an improved health status or to an increased survival probability. But, unfortunately, over the years the introduction and application of new technologies has also led to a steady expansion of the budget necessary to maintain the ever increasing level of health care in most Western countries. Therefore, concern about costs nowadays plays an increasing role in many decisions surrounding the use and implementation of medical interventions.[1,2] Whereas in the past technologies were introduced liberally, or after an assessment of merely efficacy or effectiveness, nowadays often more extensive research is done into all consequences, also in terms of costs, before a technology is implemented on a wide scale.

Medical technology assessment (MTA), which is a process of policy research in which the short and long term consequences of individual medical technologies are examined,[3] can help in making decisions on the implementation of medical technologies. The information produced by such assessment can be used by policy makers in formulating regulations and legislation, by industry in developing products, by health professionals in treating and serving patients, and by health consumers in making personal health decisions. Such technology assessments may contain information on, amongst others, the ethical and legal consequences of the new technology. But the core of almost every MTA is formed by a cost-effectiveness analysis (CE-analysis), in which the costs and effects of the technology are weighed against those of the best alternative treatment.

For the assessment of the costs and effects of medical technologies, relative to the best alternative treatment, double blind randomised controlled trials (RCTs)

in which data on both costs and effects are collected, are often advocated.[4-8] However, performing a RCT is not always possible, ethical, or feasible:

- In the evaluation of surgical techniques randomisation can be performed, but blinding is, of course, impossible.
- In the evaluation of techniques which may be life saving - like liver transplantation or neonatal extracorporeal membranous oxygenation (ECMO) - randomisation is soon thought to be unethical.[9]
- And, performing a randomised controlled double blind trial for every minor change in technology is time consuming and costly, and therefore not always feasible.

In addition to this, it should be kept in mind that RCTs may support cost-efficacy analysis rather than cost-effectiveness analysis, because of the strict inclusion criteria that are frequently applied.[5,8]

In this thesis the role modelling may have in MTA is discussed and illustrated with examples. A model of a system or process can be defined, in general, as a theoretical description that can help to understand how the system or process works, or how it might work.[10] In the following chapters *mathematical* modelling is used as a method to combine information from different sources into one coherent entity. There are many mathematical modelling techniques available, amongst which decision tree analysis, Markov processes and microsimulation are the most popular. Especially in the absence of a RCT modelling may be an efficient way to improve knowledge on a medical technology. It may also help to support the implementation of the results from RCTs or to extrapolate the results of a RCT to a more extended period of time or to a country with a different health care system.[11] And, sometimes modelling may simply represent the best option because it may lead to reliable answers much faster than any RCT ever could.

Whereas medical technology assessment nowadays frequently runs parallel with the introduction phases of a new technology, so does modelling. Modelling can be useful:

- in the development of a technology, in order to evaluate whether introduction would be necessary and/or feasible,
- in the evaluation phase of the technology, alongside its introduction, to assess efficacy and effectiveness, and
- in the implementation phase, e.g. to assess the number of treatment facilities that would be needed to provide an adequate level of care all over the country.

Table 1 gives an overview of the relation between modelling and RCTs, showing that modelling can be used in:

- the preparatory phase of a RCT, in the design of a cost-effectiveness study
- in cases where a RCT can not be fully realised; as substitution for part(s) of the trial
- to complement RCT results with cost data
- in the extrapolation of trial results and
- in the implementation of results of a trial or a cost-effectiveness study.

The application of modelling for these separate indications will be discussed below.

Modelling can be used in the *preparatory phase* of a RCT to give some impression on whether cost-effectiveness analysis would be relevant, and to make an inventory of the important issues about to come up in the final analysis. In this situation modelling is performed as part of the design of the RCT. To obtain results quickly most models from this category are based on data from literature, expert opinion and / or easily available retrospective data from medical records.
Another application of modelling in the preparatory phase lies in the development of methods needed in the subsequent RCT. If, for instance, clinical information is thought to be of importance in the analysis of the costs and effects of technologies or strategies, some kind of formalisation of this information is necessary. Modelling techniques can be used to include this information in the analysis. For example, in the analysis of the use of exercise tomographic thallium imaging in the diagnosis of severe coronary artery disease, clinical information was formalised by multivariate logistic regression,[12] and in the evaluation of costs and effects of Intensive Care facilities clinical information was combined into a model to predict mortality.[13]
And finally, modelling may be used to select the most cost-effective strategy from many possible (diagnostic) strategies. This can be done by modelling the results of a cohort study. Subsequently the most cost-effective strategy can be compared to the prevailing strategy in one prospective RCT. This approach may lead to faster results, at lower costs, than performing a large amount of RCTs, one for every strategy. With this method the most cost-effective strategy in the diagnosis of acoustic neuroma was found.[14]

In situations where RCTs are not (fully) feasible, modelling can be used to *substitute* for part(s) of the trial, or sometimes even for the whole trial.
In the evaluation of screening and prevention programmes modelling can be used to incorporate events in the future or rare - but important or expensive - events, that may be prevented by the programme, in the analysis. Models in this category often

combine data from cohort studies, epidemiological data and data from literature. Whereas trials for these programmes would require large patient populations or a very long follow-up, by modelling evaluations can be performed that would otherwise be unfeasible or too costly.
In situations were a RCT is feasible, practical and financial aspects always limit the RCT to a finite time period and number of patients. Here modelling may allow for the inclusion of rare events or events occurring in the period after completion of the trial.
In the evaluation of costs and effects of interventions in very rare diseases it is almost impossible to realise a RCT. In this situation modelling can be used to combine data from literature, and sometimes a small cohort study, to replace the whole RCT. This approach was used in the evaluation of post-exposure vaccination in rabies.[15]
And in the evaluation of technologies which may be life saving in the end stage of a disease, where no alternative treatment is available and where randomisation is thought to be unethical, like liver transplantation, modelling techniques can be used to construct a control group.[16] In this situation data from a cohort of patients treated with the new technology are used in combination with either data from literature, from historical controls or from patients from other centres or other countries.

Sometimes a RCT has already been performed to prove the efficacy of the technology under evaluation without including detailed cost analysis. If subsequently a cost-effectiveness analysis is to be performed, modelling can be used to *complement* the results of such trials with cost data.

Modelling can also be applied to *extrapolate* the trial results in order to enable a reliable assessment of cost-effectiveness.
If more than one trial has already been performed modelling can be used to combine the results of these trials.[11] This approach is often seen in pharmacoeconomic evaluations. For example in the evaluation of the efficacy of ACE-inhibitors in the treatment of heart failure several large multinational trials have been performed in only slightly different patient populations. In a subsequent cost-effectiveness analysis the results of these trials were synthesised by modelling techniques.[17]
Trials which are only designed to prove the efficacy of a new technology often use outcomes, like gastric ulcers prevented, that need translation into outcomes that are useful for CE-analysis, like QALYs or life years saved.[8,11,18,19] Modelling, based on epidemiological data or data from literature, can be used for this translation.

Often trials use a time horizon which is too short for an appropriate estimation of cost-effectiveness.[8,11,18,20] Modelling, incorporating, again, epidemiological data and data from literature, can be used to perform the extrapolation from fixed observation time to life-time costs and effects.
In addition to the extrapolations mentioned above, modelling can be applied to transform a cost-efficacy analysis into a cost-effectiveness analysis,[7,11,18,20] and to extrapolate results across indication groups or across trial settings or countries.[18]

Table 1 The relation between modelling and RCTs

Preparatory study
Modelling to identify important variables to be considered in the prospective cost-effectiveness study, and to get an impression of the outcomes to be expected
Method development, e.g. modelling to optimise the information gained from clinical observations,
Evaluation of strategies in order to select the strategies to be compared in the RCT
Substitution for (parts of) trials that can not be performed because they are thought to be unethical, unfeasible or impossible
Cost-effectiveness analysis of screening and prevention programmes focused on prohibiting rare events or events in the far future
Incorporation of relatively rare events, or events occurring in the period after completion of the trial
Cost-effectiveness analysis for diseases with a low prevalence or incidence
Modelling of the control group
Complementation of trial results with cost data
Extrapolation of trial results
Meta-analysis of results of similar clinical trials to enable estimation of a reliable cost-effectiveness ratio
Extrapolation of intermediate trial outcomes to final outcomes for cost-effectiveness analysis (in terms of QALYs or life years gained)
Extrapolation of trial results to a longer period
Transformation of a cost-efficacy analysis into a cost-effectiveness analysis
Extrapolation of trial results to countries with a different health care system, different treatment patterns and/or different unit prices
Implementation of the results of a trial or a cost-effectiveness analysis
Macro-economic consequences of the introduction of a new technology
Analysis of the need for (additional) treatment facilities, resource allocation

Modelling can also be used to support the actual *implementation* of a new technology. Here decision makers are interested, not only in information about the cost-effectiveness of a new technology, but also in the macroeconomic consequences of the introduction of the technology. Modelling techniques, incorporating information from various sources, can help to make an estimation of the consequences of introducing a technology on a nation-wide scale.
And, finally, modelling techniques derived from operations research can be used in resource allocation, e.g. to evaluate the consequences of the introduction of additional treatment facilities.[21-27]

Table 1 shows that modelling may have a place in many different phases of a medical technology assessment. This table aims at giving an overview of the many possible applications in modelling in relation to RCTs. However, the categories should not be taken too strict and are not always mutually exclusive. Many cost-effectiveness analyses can be ascribed to more than one category. For example, if the results of a trial are extrapolated, often more than one extrapolation is done. And sometimes the boundaries between categories are unclear, for example the difference between a preparatory study based on a large number of trials published in literature and an extrapolation of trial results through meta-analysis of a limited number of similar trials often is only gradual.

That modelling in theory is already closely associated with the concept of cost-effectiveness analysis is illustrated by the fact that the author of a recent editorial on this subject, that was published in the New England Journal, assumed that modelling techniques were applied in <u>all</u> cost-effectiveness analyses.[1,28,29] However, in practice modelled cost-effectiveness analysis still only comprises a very small part of medical literature. A MEDLINE survey shows that only 5.3% of the more than 350,000 registered publications for 1994 mention the word "model" (or modelling) in either title, abstract or subject heading, 0.5% the word "costs", and 0.1% the word "cost-effectiveness". Only 135 publications mention both "costs" and "model" (0.04%) and only 57 the words "cost-effectiveness" and "model" (0.02%). Thirteen of these 57 articles use the word model in an other context or do not actually present a cost-effectiveness analysis, three were review articles. Table 2 gives an overview of the remaining 41 articles. Modelling techniques seem to be especially popular for cost-effectiveness analysis in screening, prevention or vaccination programmes. Most of the other studies are preparatory studies, designed to give a "quick and dirty" impression of the costs and effects that are to be expected after introduction of a new technology. However, it looks like only few

of these studies will actually be followed by a RCT. There were no studies among the 41 that described the incorporation of rare events in a RCT, or a cohort study with a modelled control group. None of the studies was devised to convert a cost-efficacy analysis into a cost-effectiveness analysis. Decision tree analysis appears to be the most popular modelling technique, featuring in 20 of the analyses, probably because of the general availability of decision analytic software.

As stated before, the purpose of this thesis is to illustrate the use of modelling techniques in different phases of a medical technology assessment, and therefore also in different phases of the introduction of a medical technology, to discuss the relation between modelling and RCTs, and to illustrate some of the many different techniques that can be used in modelling.

In chapter 1 an example is given of a modelling approach that was used as a preparatory study for a randomised controlled trial. In 1993, parties in the health care sector in the Netherlands were discussing the introduction of helicopter medical crews for the treatment of seriously wounded trauma victims. In this discussion both supporters and adversaries being so convinced of their own point of view deemed a randomised controlled trial to study the cost-effectiveness of this new facility as unnecessary. They assumed the facility evidently would - respectively wouldn't - be cost-effective. To shed more light on this rather dark field, a "quick and dirty" modelling approach was chosen to evaluate whether the facility would be cost-effective and whether a RCT would be necessary.

Chapter 2 gives an example of a modelling technique used to formalise clinical information in patients with suspected pulmonary embolism. Especially in pulmonary embolism many unspecific clinical items - like cough, shortness of breath, tachycardia - raise the suspicion of the presence of pulmonary embolism. It was thought that structuring this information into a clinical decision rule could lead to cost-savings and an improved diagnosis.
Subsequently, the optimal place of this clinical decision rule within the diagnostic strategy was sought in a cost-effectiveness analysis, which is presented in chapter 3. In this chapter the use of a decision model in the evaluation of the costs and effects of several diagnostic strategies is illustrated. The model permits the evaluation and comparison of the costs and effects of an almost unlimited number of diagnostic strategies.

Table 2 *Overview of the 41 articles on cost-effectiveness analysis and modelling published in 1994 derived from MEDLINE. compl. = complementation; d = deterministic; dt = decision tree; impl. = implementation; l = logistic regression; LISREL = Linear Structural RELations; m = Markov process; mfc = multicompartment with fuzzy control; ns = not specified; s = simulation. * = only abstract available.*

	preparatory			substitution				compl.	extrapolation					impl.		
	model instead of experimental design	method development	evaluation of strategies	screening / prevention	incorporation rare events / events after trial period	diseases with low prevalence or incidence	control group modelled	complementation of trial results with cost data	extrapolation intermediate outcomes RCT	extrapolation RCT to a longer period	combining results similar RCTs	translation efficacy - effectiveness	extrapolation RCT / CE analysis to other countries	macro-economic consequences	resource allocation	modelling technique
Arikian[30]	x												x			dt
AuBuchon[31]				x										x		dt,m
Bance[14]			x													dt
Bellazi[32]		x														mfc,s
Brown[33]								x		x				x		s
Busch[34]									x							LISREL
Cantor[15]				x		x										dt
Christian[12]		x	x													l
Cohen[35]	x															dt,m
Davey[36]	x															ns
Ebell[37]	x															dt
Einarson[38]	x															ns
Geelhoed[39]				x												dt,m
Goodnough[40]				x												m,dt
Healy[41]				x												dt
Iinuma*[42]				x												ns
Johnson[43]								x		x	x		x	x		dt
Katz[44]	x															dt,m

	preparatory		substitution		compl.	extrapolation				impl.	
	model instead of experimental design	method development	evaluation of strategies	screening / prevention	incorporation rare events / events after trial period	diseases with low prevalence or incidence	control group modelled	complementation of trial results with cost data	extrapolation intermediate outcomes RCT	extrapolation RCT to a longer period	combining results similar RCTs	translation efficacy - effectiveness	extrapolation RCT / CE analysis to other countries	macro-economic consequences	resource allocation	modelling technique
Kwan-Gett[45]	x															dt
Lacour[46]				x										x		s
Lenfant[47]				x										x		ns
Lessler[48]	x															dt
Lieu[49]				x										x		dt
Lindfors[50]				x												m
Lurie[51]				x										x		dt
Martens[52]				x					x							ns
McFarland[53]	x															ns
McIntyre[54]				x												dt
Michel[55]	x															d
Murray[56]															x	ns
Oster[57]				x				x	x	x						dt,m
Paladino[58]	x															dt
Paul[17]								x		x	x					m,dt
Phatak[59]				x												dt
Rapoport[60]		x														ns
Rosenquist[61]				x												m
Rowley[62]				x												d,s
Sell[63]				x										x		ns
Shiell[64]	x															m
Walan[65]								x		x	x			x		s
Watcha[66]	x															dt
Total	**13**	**3**	**2**	**18**	**0**	**1**	**0**	**5**	**3**	**5**	**3**	**0**	**3**	**8**	**1**	

In chapter 4 the use of a modelling technique to "make-up" for the absence of a control group is illustrated with an example in liver transplantation. In the evaluation of the cost-effectiveness of liver transplantation in the Netherlands a RCT was thought to be unethical, because in the USA it was already accepted (but not proven!) that transplantation could lead to an improved survival probability in patients for whom little alternative treatment was available. In the Dutch cost-effectiveness study an adapted design was chosen. Liver transplantation was introduced on a limited scale and both costs and survival were monitored. To prove the effectiveness of the transplantation a Cox regression model was used to predict survival as it would have been without transplantation. In chapter 4A the technique, its limits and alternatives are discussed more in detail. In chapter 4B the use of the technique in the cost-effectiveness analysis of liver transplantation is illustrated.

Chapter 5 illustrates the use of modelling techniques in the complementation and extrapolation of trial results. Here a Markov model is used to extrapolate the results of a large multi-centre North-American trial, on the use of ACE-inhibitors for patients with asymptomatic left ventricular dysfunction, to the Netherlands.[67] The trial results were complemented with cost data specific for the Dutch situation. Also, an extrapolation of the results to a longer treatment period was made (20 years instead of 4 years). Part of the translation from cost-efficacy analysis to cost-effectiveness analysis was made by introducing the Dutch treatment policy in clinical practice as the best alternative treatment instead of the placebo treatment used in the original trial.
In this chapter also the use of modelling techniques for the evaluation of the macroeconomic consequences of the introduction of ACE-inhibitors as preventive therapy is shown.

Finally, chapter 6 illustrates the use of a model in the planning of neonatal ECMO facilities in the Netherlands. Instead of introducing treatment facilities on a trial and error basis, microsimulation was used to predict the number of patients that would have to be referred to centres abroad, relative to the number of treatment facilities and the number of neonates requiring treatment annually.

References

1 Kassirer JP, Angell M. The Journal's policy on cost-effectiveness analyses [editorial]. New England Journal of Medicine 1994; 331(10): 669-670

2 Robinson R. Economic evaluation and health care. The policy context. British Medical Journal 1993; 307: 994-996

3 Office of Technology Assessment. Strategies for Medical Technology Assessment. Washington: Office of Technology Assessment, 1982

4 Robinson R. Economic evaluation and health care. What does it mean? British Medical Journal 1993; 307: 670-3

5 Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized controlled trials. Medical Care 1992; 30(3): 231-43

6 Drummond MF, Davies L. Economic analysis alongside clinical trials. Revisiting the methodological issues. International Journal of Technology Assessment in Health Care 1991; 7(4): 561-573

7 Drummond MF, Ludbrook A, Lowson K, Steele A. Studies in economic appraisal in health care. Volume 2. Oxford: Oxford University Press, 1986

8 Sloan FA (editor). Valuing health care. Costs, benefits, and effectiveness of pharmaceuticals and other medical technologies. Cambridge: Cambridge University Press, 1995

9 Hilhorst, MT. The ethical assessment of new technologies: the case of ECMO. In: Misbin RI, Jennings B et al. Health care crisis? The search for answers. Frederick, Maryland: Univ Publ Group, 1995; 200-209

10 Collins COBUILD English language dictionary. London: HarperCollins Publishers, 1994

11 Langley PC. Outcomes research and modeling therapeutic interventions for economic evaluations. Clin Ther 1994; 16(3): 538-52

12 Christian TF, Miller TD, Bailey KR, Gibbons RJ. Exercise tomographic thallium-201 imaging in patients with severe coronary artery disease and normal electrocardiograms. Annals of Internal Medicine 1994; 121(11): 825-32

13 Rapoport J, Teres D, Lemeshow S, Gehlbach S. A method for assessing the clinical performance and cost-effectiveness of intensive care units: a multicenter inception cohort study. Critical Care Medicine 1994; 22(9): 1385-91.

14 Bance ML, Hyde ML, Malizia K. Decision and cost analysis in acoustic neuroma diagnosis. Journal of Otolaryngology 1994; 23(2):109-20.

15 Cantor SB, Clover RD, Thompson RF. A decision-analytic approach to postexposure rabies prophylaxis. American Journal of Public Health 1994; 84(7): 1144-8.

16 Bonsel GJ, Klompmaker IJ, van't Veer F, et al. Use of prognostic models for assessment of value of liver transplantation in primary biliary cirrhosis. Lancet 1990; 335: 493-497

17 Paul SD, Kuntz KM, Eagle KA, Weinstein MC. Costs and effectiveness of angiotensin converting enzyme inhibition in patients with congestive heart failure. Archives of Internal Medicine 1994; 154(10): 1143-9.

18 Commonwealth of Australia. Draft guidelines for the pharmaceutical industry on preparation of submissions to the Pharmaceutical Benefits Committee. Appendix J. Canberra: department of Health, Housing and Community services, 1995.

19 Gabriel SE, Campion ME, O'Fallon WM. A cost-utility analysis of misoprostol prophylaxis for rheumatoid arthritis patients receiving nonsteroidal antiinflammatory drugs. Arthritis & Rheumatism 1994; 37: 333-341

20 Canadian Coordinating Office for Health Technology Assessment. Guidelines for economic evaluation of pharmaceuticals: Canada. 1st edition. Ottawa, CCOHTA, 1994

21 Bolger PG, Davies R. Simulation model for planning renal services in a district health authority. BMJ 1192; 305: 605-608

22 Davies H, Davies R. A simulation model for planning services for renal patients in Europe. J Opl Res Soc 1987; 38(8): 693-700

23 Davies R, Davies T. Using simulation to plan health services resources: discussion paper. J Roy Soc Med 1986; 79: 154-157

24 Dumas MB. Hospital bed utilization: an implemented simulation approach to adjusting and maintaining appropriate levels. Health services research 1985; 20(1): 43-61

25 Feest TG, Harrison P. Simulation model for planning renal services (letter). BMJ 1992;305:1018

26 Harris RA. Hospital bed requirements planning. European Journal of Operational Research 1985; 25: 121-126

27 Romanin-Jacur G, Facchin P. Optimal planning of a pediatric semi-intensive care unit via simulation. European Journal of Operational Research 1987; 29: 192-198

28 Steinberg EP. Cost-effectiveness analysis [letter]. New England Journal of Medicine 1995; 332(2): 123

29 Schulman K. Cost-effectiveness analysis [letter]. New England Journal of Medicine 1995; 332(2): 124

30 Arikian SR, Einarson TR, Kobelt-Nguyen G, Schubert F. A multinational pharmaco-economic analysis of oral therapies for onychomycosis. The Onychomycosis Study Group. British Journal of Dermatology 1994; 130 Suppl 43: 35-44

31 AuBuchon JP, Birkmeyer JD. Safety and cost-effectiveness of solvent-detergent-treated plasma. In search of a zero-risk blood supply. JAMA 1994 272(15): 1210-4.

32 Bellazzi R, Siviero C, Bellazzi R. Mathematical modeling of erythropoietin therapy in uremic anemia. Does it improve cost-effectiveness? Haematologica 1994; 79(2): 154-64.

33 Brown ML, Nayfield SG, Shibley LM. Adjuvant therapy for stage III colon cancer: economics returns to research and cost-effectiveness of treatment. Journal of the National Cancer Institute 1994; 86(6): 424-30

34 Busch P, Schwendener P, Leu RE, von Dach B, Castiglione M. Life quality assessment of breast cancer patients receiving adjuvant therapy using incomplete data. Health Economics 1994; 3(4): 213-20.

35 Cohen DJ, Breall JA, Ho KK, Kuntz RE, Goldman L, Baim DS, Weinstein MC. Evaluating the potential cost-effectiveness of stenting as a treatment for symptomatic single-vessel coronary disease. Use of a decision-analytic model. Circulation 1994; 89(4): 1859-74.

36 Davey P, Rutherford D, Graham B, Lynch B, Malek M. Repeat consultations after antibiotic prescribing for respiratory infection: a study in one general practice. British Journal of General Practice 1994; 44(388): 509-13.

37 Ebell MH, Kruse JA. A proposed model for the cost of cardiopulmonary resuscitation. Medical Care1994; 32(6): 640-9.

38 Einarson TR, Arikian SR, Shear NH. Cost-effectiveness analysis for onychomycosis therapy in Canada from a government perspective. British Journal of Dermatology 1994; 130 Suppl 43: 32-4.

39 Geelhoed E, Harris A, Prince R. Cost-effectiveness analysis of hormone replacement therapy and lifestyle intervention for hip fracture. Australian Journal of Public Health 1994; 18(2):153-60.

40 Goodnough LM, Grishaber JE, Birkmeyer JD, Monk TG, Catalona WJ. Efficacy and cost-effectiveness of autologous blood predeposit in patients undergoing radical prostatectomy procedures. Urology 1994; 44(2): 226-31.

41 Healy JC, Frankforter SA, Graves BK, Reddy RL, Beck JR. Preoperative autologous blood donation in total-hip arthroplasty. A cost-effectiveness analysis. Archives of Pathology & Laboratory Medicine 1994; 118(4): 465-70

42 Iinuma T, Tateno Y, Matsumoto T. Comparison of two types of mass screening for lung cancer in terms of cost-effectiveness: indirect chest X-ray vs LSCT. [in Japanese] Nippon Igaku Hoshasen Gakkai Zasshi - Nippon Acta Radiologica 1994; 54(10): 943-9.

43 Jonsson B, Bebbington PE. What price depression? The cost of depression and the cost-effectiveness of pharmacological treatment. British Journal of Psychiatry 1994; 164(5): 665-73.

44 Katz DA, Cronenwett JL. The cost-effectiveness of early surgery versus watchful waiting in the management of small abdominal aortic aneurysms. Journal of Vascular Surgery 1994; 19(6): 980-90.

45 Kwan-Gett TS, Whitaker RC, Kemper KJ. A cost-effectiveness analysis of prevaccination testing for hepatitis B in adolescents and preadolescents. Archives of Pediatrics & Adolescent Medicine 1994; 148(9): 915-20.

46 Lacour A, Chauvin F, Duru G, Philip T. Economic evaluation of screening for neuroblastoma. [in French] Revue d'Epidemiologie et de Santé Publique 1994; 42(1): 24-33.

47 Lenfant C, Roccella EJ. Preventing strokes: considerations for developing health policy. Health Reports1994; 6(1): 216-23.

48 Lessler DS, Sullivan SD. Stergachis A. Cost-effectiveness of unenhanced MR imaging vs contrast-enhanced CT of the abdomen or pelvis. American Journal of Roentgenology 1994; 163(1): 5-9.

49 Lieu TA, Cochi SL, Black SB, Halloran ME, Shinefield HR, Holmes SJ, Wharton M, Washington AE. Cost-effectiveness of a routine varicella vaccination program for US children. JAMA 1994; 271(5): 375-81.

50 Lindfors KK, Rosenquist CJ. Needle core biopsy guided with mammography: a study of cost-effectiveness. Radiology 1994; 190(1): 217-22.

51 Lurie P, Avins AL, Phillips KA, Kahn JG, Lowe RA, Ciccarone D. The cost-effectiveness of voluntary counseling and testing of hospital inpatients for HIV infection. JAMA 1994; 272(23): 1832-8.

52 Martens LL, Guibert R. Cost-effectiveness analysis of lipid-modifying therapy in Canada: comparison of HMG-CoA reductase inhibitors in the primary prevention of coronary heart disease. Clinical Therapeutics 1994; 16(6): 1052-62

53 McFarland BH. Cost-effectiveness considerations for managed care systems: treating depression in primary care. American Journal of Medicine 1994; 97(6A): 47S-57S.

54 McIntyre P, Hall J, Leeder S. An economic analysis of alternatives for childhood immunisation against Haemophilus influenzae type b disease. Australian Journal of Public Health 1994; 18(4): 394-400.

55 Michel BC, van Hout BA. Is use of helicopters with medical crews in severely injured accident victims cost-effective in The Netherlands? A modelling approach. [in Dutch]. Nederlands Tijdschrift voor Geneeskunde 1994; 138(46): 2310-5.

56 Murray CJ, Kreuser J, Whang W. Cost-effectiveness analysis and policy choices: investing in health systems. Bulletin of the World Health Organization 1994; 72(4): 663-74.

57 Oster G, Huse DM, Lacey MJ, Epstein AM. Cost-effectiveness of ticlopidine in preventing stroke in high-risk patients. Stroke 1994; 25(6): 1149-56.

58 Paladino JA, Fell RE. Pharmacoeconomic analysis of cefmenoxime dual individualization in the treatment of nosocomial pneumonia. Annals of Pharmacotherapy 1994; 28(3): 384-9.

59 Phatak PD, Guzman G, Woll JE, Robeson A, Phelps CE. Cost-effectiveness of screening for hereditary hemochromatosis. Archives of Internal Medicine 1994; 154(7): 769-76.

60 Rosenquist CJ, Lindfors KK. Screening mammography in women aged 40-49 years: analysis of cost-effectiveness. Radiology 1994; 191(3): 647-50.

61 Rowley JT, Anderson RM. Modelling the impact and cost-effectiveness of HIV prevention efforts. AIDS 1994; 8(4): 539-48.

62 Ruebush TK, Zeissig R, Koplan JP, Klein RE, Godoy HA. Community participation in malaria surveillance and treatment. III. An evaluation of modifications in the Volunteer Collaborator Network of Guatemala. American Journal of Tropical Medicine & Hygiene 1994; 50(1): 85-98.

63 Sell RL, Jovell AJ, Siegel JE. HIV screening of surgeons and dentists: a cost-effectiveness analysis. Infection Control & Hospital Epidemiology 1994; 15(10): 635-45.

64 Shiell A, Briggs A, Farrell GC. The cost-effectiveness of alpha interferon in the treatment of chronic active hepatitis C. Medical Journal of Australia 1994; 160(5): 268-72.

65 Walan A, Eriksson S. Long-term consequences with regard to clinical outcome and cost-effectiveness of episodic treatment with omeprazole or ranitidine for healing of duodenal ulcer. Scandinavian Journal of Gastroenterology 1994; Supplement 201: 91-7.

66 Watcha MF, Smith I. Cost-effectiveness analysis of antiemetic therapy for ambulatory surgery. Journal of Clinical Anesthesia 1994; 6(5): 370-7.

66 Pfeffer MA, Braunwald E, Moyé LA, et al. Effect of captopril on mortality and morbidity in patients with left ventricular dysfunction after myocardial infarction. Results of the Survival and Ventricular Enlargement trial. New England Journal of Medicine 1992; 327: 669-677

IS USE OF HELICOPTERS WITH MEDICAL CREWS FOR SERIOUSLY INJURED ACCIDENT VICTIMS COST-EFFECTIVE IN THE NETHERLANDS? A MODELLING APPROACH*

Summary

In this chapter the cost-effectiveness of helicopter medical crews for patients with major trauma in the Dutch situation is assessed. For this assessment a modelling approach is used, combining available information with plausible assumptions. First, a base-line scenario is developed, implementing information that is available in the literature and a few assumptions. Next, the sensitivity of the cost-effectiveness for the various assumptions is evaluated in alternative scenarios. For the separate scenarios the relation between the costs per life year gained and the expected increase in the number of patients surviving an accident is established. In the base-line scenario the number of victims with a trauma score ≤ 12 is estimated at 4500 per year, it is assumed that victims who die at the scene of the accident are unsalvageable and that hospital treatment costs DFL 27,000 per patient. As a result costs per life year saved are only under DFL 50,000 when more than 29 patients per year (9.23% gain in survival) are additionally saved. Sensitivity analysis shows that the cost-effectiveness depends highly on the expected gain in survival and on the estimate of the costs of hospital treatment. Both variables are surrounded with a lot of uncertainty. The model shows that, relying on the information that is available at this moment, no definite conclusions can be drawn on the cost-effectiveness of this facility, because too many essential data are missing. Empirical research is necessary.

* Based on: Michel BC, Hout BA van. Is het inzetten van helikopters met een traumateam bij ernstig gewonde ongevalsslachtoffers in Nederland kosten-effectief? Een modelmatige benadering. Nederlands Tijdschrift voor Geneeskunde 1994; 138(46): 2310-2315

Introduction

Because the first hours after the accident are of vital interest to seriously injured accident victims, means to improve the medical assistance in this period are always searched for. In Germany the start of treatment for accident victims was shifted towards the scene of the accident several years ago. There, concurrently with an ambulance staffed by paramedics, a physician is sent to the scene of the accident by helicopter.[1-5] The physician is charged with stabilising the patient. He (or she) also decides to what hospital the patient will be sent, and by which means of transport. In this system the helicopter is not primarily intended for the transportation of the patient, but mainly for the transportation of the physician. However, helicopter transportation of the patient is possible.

THE PROBLEM

Would helicopters with a medical crew be a welcome improvement to the current system in the Dutch situation? On the one hand there are some trauma specialists, supported by the ANWB, who answer this question affirmatively;[6] on the other hand there is a letter from the former Secretary of State Simons to the Dutch Parliament in which the introduction of helicopters is called unnecessary.[7] Ideally the answer to this question should depend on the balance between costs and effects,[5,8] and for a reliable assessment of this cost-effectiveness ratio empirical data are necessary. However, both supporters and adversaries of helicopters with medical crews implicitly assume that empirical research is superfluous, because they regard their own point of view as evident. In the modelling approach presented here, it is evaluated whether such peremptory statements can really be made, if one is to rely on the information that is currently available for the Dutch situation.

THE ADDITIONAL VALUE OF A HELICOPTER WITH A MEDICAL CREW TO THE CURRENT SYSTEM

A helicopter can supply medical assistance within a radius of 70 kilometres just as fast as an ambulance can in a radius of 15 kilometres. By helicopter a limited number of "trauma teams" can be employed to cover the whole of the Netherlands.

Patients can be transported to specialised centres over larger distances and distributed over hospitals with more expertise,[9] thereby probably diminishing the number of secondary transports. Also, it is likely that better adaptations for shortages in Intensive Care facilities in the region of the accident can be made.

An additional advantage might be found in the number of trauma centres in the Netherlands. It is almost certain that these centres will eventually be installed, but the number, time-line and form are still under discussion.[4,5,7,10,11] If helicopters are stationed at trauma centres, probably less than the 11 centres proposed by the College for Hospital Facilities (CHF) will be needed, because by helicopter patients can be transported to the centres over a larger distance.[4]

Estimates

In the estimates of costs and effects, in terms of Dutch Guilders and life years gained, data specific for the Netherlands are used whenever possible. If no data are available assumptions are made. All estimates are combined into a descriptive model in a computer spreadsheet programme (Quattro Pro, version 5.00, Borland International Inc.). In this model first the costs and effects are calculated for a baseline scenario. Subsequently the influence of the assumptions on the results is evaluated in alternative scenarios.

ESTIMATE OF THE NUMBER OF ANNUAL FLIGHTS

International publications show that patients with a trauma score less than or equal to 12 have a bad prognosis and might benefit from care in a specialised centre, so maybe also from transportation by helicopter.[4,12,13,14] Therefore, the model is limited to these patients. Little data are available on the exact number of patients within this category. The CHF combines three separate estimates into a final estimate of 4500 seriously wounded trauma patients admitted into hospital each year.[4]

To estimate the annual number of helicopter flights, accidents in which victims die at the scene of the accident should also be taken into account. According to the Central Bureau for Statistics 1280 persons died within 30 days of a *traffic* accident

in 1991, whereas 698 of these victims died at the scene of the accident.[15] Publications by Draaisma state that the ratio of traffic accidents to other accidents is 81.4% to 18.6% in seriously injured patients.[17] If the assumption is made that this ratio is identical for victims who die at the scene of the accident, an estimate of 4500 + (698/0.814) = 5357 seriously or deadly injured accident victims results.

It is assumed that the helicopter can be used only during daytime, in reasonable weather conditions, and outside city areas. The Central Bureau for Statistics reports that 29.5% of traffic accidents in which people are killed or transported to hospital take place during daytime outside the city area and that weather conditions are favourable for helicopter usage in 97.2% of the traffic accidents.[15] Since corresponding figures for non-traffic accidents are unknown, it is assumed that the same percentages will apply. The potential number of accident victims for whom a helicopter can be used can therefore be estimated at 5357 x 0.295 x 0.972 = 1537. Two hundred and forty six of these victims will die at the scene of the accident, before transportation to a hospital can take place.

In practice, the helicopter is used rather frequently for patients who do not really need this kind of assistance. Data from the Austrian and German helicopter services show that respectively 7% (200/2874) en 25.2% (7893/31251) of their flights were unnecessary.[16,18] In the model the average of these two percentages is included: 16.1%.

If it is assumed that on average one seriously or deadly injured patient will be involved in every accident for which the helicopter is employed, it can be estimated that 1784 annual helicopter flights will be made.

ESTIMATE OF THE NUMBER OF LIFE YEARS GAINED

The Dutch Central Bureau for Statistics publishes data on the age and sex of deceased traffic accident victims.[15]
Because death may only be declared after a personal inspection by a physician, who currently is not present at the scene of the accident,[19] it is plausible that victims registered as "deceased at the scene of the accident" can't be saved even by a helicopter with a medical crew. The Central Bureau for Statistics' accident statistics show that yearly 582 persons die within 30 days after a traffic accident, if deaths at the scene of the accident are excluded.[15] After correction for the time of

the accident - 73% of the fatal accidents take place in daylight or twilight -, the place of the accident - 60.6% outside the city area -, weather conditions - in 98% the weather conditions allow a helicopter flight -, and non-traffic accident victims, an estimate results of (582 x 0.73 x 0.606 x 0.98) / 0.814 = 310 deaths that might be prevented by a helicopter service. By relating these data to the average expected life-time for age category and sex, it can be estimated that 12641 life years may be saved by the helicopter service, 40.8 per deceased.[20]

ESTIMATE OF THE COSTS

The estimate of the costs of the helicopter service is based on a research project performed by the Rotterdam School of Management in 1987.[21] The fixed costs per helicopter are estimated at DFL 2,030,000 annually (table I). The variable costs are estimated at DFL 750 per flight, for maintenance, fuel and (spare) parts.

Table 1 Fixed costs per helicopter

		Annual costs
Rent helicopter	DFL	450,000
Insurance helicopter (6% of its value)	DFL	250,000
Insurance medical personnel	DFL	100,000
Communication-assistant	DFL	50,000
Training personnel	DFL	50,000
Hangar	DFL	30,000
Maintenance and other	DFL	100,000
Organisation	DFL	100,000
Communication devices	DFL	50,000
Two pilots (average salary DFL 125,000)	DFL	250,000
Medical personnel[a]	DFL	600,000
Total	**DFL**	**2,030,000**

[a] Two physicians (average salary DFL 200,000) and two nurses (average salary DFL 100,000)

Because part of the effect of the use of helicopters is due to the treatment of patients in trauma centres instead of general hospitals, the costs of these centres are included in the calculations, but only as far as they can be attributed to patients in which the helicopter-service is involved. The CHF estimates that the establishment of 11 trauma centres will cost DFL 100,000,000 each year to acquire a fixed gain in survival of 25%.[4] In addition to the calculations of the CHF changes in costs due to changes in patient distribution are taken into account here. And it is assumed that, after the introduction of the helicopters, only five trauma centres need to be established (one per helicopter), and that for each of these five centres no more capital investments will be necessary than in a situation with eleven centres. The costs of hospital treatment for an accident victim are estimated by the CHF at DFL 27,000 to DFL 48,000. In the baseline scenario the lowest of these two amounts is used.

In agreement with the assumptions of the CHF costs of treatment in general hospitals are estimated at 55% of the costs of treatment in trauma centres, and the costs of treatment for a patient who eventually dies at half of normal treatment costs. Also in accordance with the CHF, it is assumed that in the current situation already 25% of the patients are treated in hospitals comparable to trauma centres. Table 2 gives an overview of the separate scenarios.

Table 2 *Overview of the separate scenarios. The limit of cost-effectiveness is DFL 50,000 for all scenarios, there are 5 helicopters and the results are discounted by 5%. no. = number of patients; TS = trauma score; avg = average; late mortality = mortality more than 30 days after the accident*

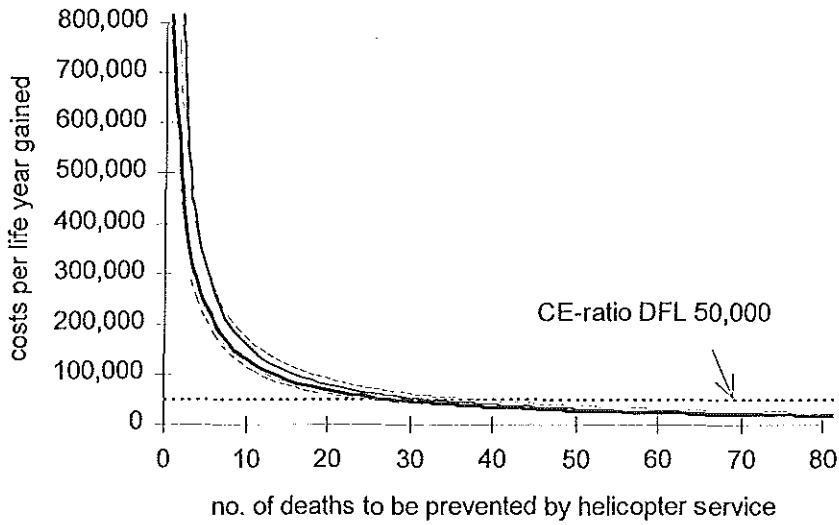
Scenario	avg. costs hospital treatment	can deceased at scene accident be saved?	no. with TS ≤ 12 annually	life expectancy relative to age and sex	late mortality
Baseline	*f* 27.000,-	no	4500	100%	0%
Scenario I	*f* 48.000,-	no	4500	100%	0%
Scenario II	*f* 27.000,-	yes	4500	100%	0%
Scenario III	*f* 27.000,-	no	3500	100%	0%
Scenario IV	*f* 27.000,-	no	4500	75%	0%
Scenario V	*f* 27.000,-	no	4500	100%	10%

Results

The cost-effectiveness of a helicopter service for seriously injured trauma patients is expressed in additional costs per life year gained compared to the current situation. Both costs and life years gained are discounted by 5% annually, according to the guidelines for cost calculations in cost-effectiveness analysis.[21] This means that costs and effects in the far future are given less weight than costs and effects in the near future.

For the Netherlands, a nation-wide service can probably be provided with five helicopters.[22] Total fixed costs will amount to DFL 10,150,000. Variable costs will depend on the number of flights, and will amount to approximately DFL 1,338,000, if 1784 annual flights are performed. To these costs, costs for treatment in specialised centres, partly depending on the percentage of patients surviving, are added.

Figure 1 *Number of deceased to be saved in the separate scenarios in order to attain a certain cost-effectiveness ratio.(— baseline; – – – = scenario I; – - – - – = scenario II; - - - - = scenario III; —— = scenario IV; – - – - – - = scenario V)*



If all deaths could be prevented by the helicopter service the cost-effectiveness would be most favourable: DFL 5.074,- per life year gained. However, it is reasonable to assume that part of the mortality after accidents can

never be prevented. Figure 1 shows the cost-effectiveness in Dutch Guilders per life year saved for different numbers of deaths prevented by the helicopter service, as compared to the current situation. The limit of DFL 50,000 per life year saved, that was considered cost-effective in, for example, kidney-function-replacing therapy, and heart and liver transplantation,[23-25] is reached if the helicopter service will be life-saving in 29 persons annually. Table 3 gives a summary of the results and a comparison to the current situation.

Table 3 *Costs in the baseline scenario (gain in survival 9.23%) compared to the current situation in millions Dutch Guilders. (no. = number of patients)*

	current situation		after introduction helicopters	
	no. pat	costs	no. pat	costs
Patients in trauma centre: survivors	262	7.1	1048	
Patients in trauma centre: dead < 30 days	44	0.6	177	2.4
Capital investment		-		1.9
Total costs treatment in trauma centres		*7.7*		*32.6*
Patients in general hospital: survivors	768	11.4	-	
Patients in general hospital: dead < 30 days	151	1.1	-	
Total costs treatment in general hospital		*12.5*		-
Total costs		**20.2**		**32.6**

SENSITIVITY ANALYSIS

The influence of the simplifications and assumptions that were used in the baseline model is tested in some alternative scenarios. Differences to the baseline scenario are shown in table 2, results in table 4. In the presentation of the results a preferred cost-effectiveness of DFL 50,000 per life year gained is used.

The cost-effectiveness ratio is rather sensitive to the average costs of hospital treatment for trauma patients. In scenario I the largest of the two amounts stated by the CHF is used (DFL 48,000), in that case a much larger gain in survival must be attained to reach the limit for cost-effectiveness of DFL 50,000.[4] In scenario II it is

assumed that victims who die at the scene of the accident can also be saved by the helicopter service. This does not result in a larger number of flights, but it does result in a larger number of hospital admissions.

The results of scenario III show that changes in the number of patients with a trauma score equal to or smaller than 12 only have a minor influence on the outcome. However, the number of additional treatments in a trauma centre is clearly smaller, 715 a year.

In scenario IV it is assumed that life expectancy is shortened by 25%, due to remaining morbidity. In scenario V the assumption is made that in the period following 30 days after the accident, the registration period of the Central Bureau for Statistics, still 10% of the trauma patients will die and that these patients can be saved by better primary medical assistance.

Finally, the influence of the percentage of unnecessary flights on the cost-effectiveness is evaluated. This influence appears to be very minor. Without unnecessary flights, with complete efficiency of the service, in the baseline scenario the limit of DFL 50,000 is attained if 28 persons are saved, whereas with 50% unnecessary flights the limit is reached if 29 persons are saved. In this, however, no account is taken of the loss in efficiency that may occur if, by an increase in the number of flights, the number of simultaneous requests for helicopter assistance increases.
The cost-effectiveness of the use of helicopters with a trauma team appears to depend strongly on the average costs of hospital treatment, and on the gain in survival that can be reached with this service (see figure 1). Other factors, like the annual number of patients presenting with a trauma score less than or equal to 12, the life expectancy, or the percentage of flights that are unnecessary turn out to be less important.

Table 4 *Expected costs for society (in millions Dutch Guilders) attributed to the helicopter service, and minimum gain in survival necessary to obtain a cost-effectiveness ratio of DFL 50,000 per life year gained, for the separate scenarios. no. = number*

	baseline	scenario I	scenario II	scenario III	scenario IV	scenario V
Extra patients in trauma centre	919	919	951	715	919	919
Minimum gain in survival needed for cost-effectiveness	9.23%	12.29%	4.85%	7.73%	10.24%	8.35%
No. of patients to be saved / no. deceased considered salvageable	29/310	38/310	28/574	24/310	32/310	28/341
Costs for 5 trauma centres	32.6	56.2	33.5	25.7	32.6	32.3
Costs in current situation	- 20.2	- 35.9	- 20.2	- 16.9	- 20.2	- 20.0
Extra costs 5 trauma centres	12.4	20.3	13.3	8.8	12.4	12.3
Costs 5 helicopters	11.5	11.5	11.5	11.2	11.5	11.5
Total costs	**23.9**	**31.8**	**24.8**	**20.0**	**23.9**	**23.8**

Discussion

In the absence of empirical research an attempt was made to model the cost-effectiveness of the use of helicopters with medical crews using available data. In this approach several factors had to be left out of account because many data were unavailable. For instance, some patients might need shorter hospitalisation and have less remaining symptoms after helicopter assistance, due to the fast initiation of treatment.[1] However, helicopter assistance might also keep patients who would have died otherwise alive in a bad condition, causing long hospitalisations in nursing homes or rehabilitation centres.[8] The (financial) consequences of a decrease in the number of secondary transports were also left out of account. In the current analysis the indirect costs, like loss of production for society due to illness, disablement and death were not included. Conform the guidelines for cost calculations in health care research this is allowed if it is to be expected that the period of sick-leave will be longer than the "friction period" in all scenarios.[21] This friction-period represents the period between the employee dropping out and his or her replacement from inside or outside the company.[26] In the Netherlands this friction period was 96 days on average in 1990.[27] The average sick-leave for fractures due to traffic accidents was longer, on average approximately 126 days. [28] This last figure is not corrected for the level of trauma, for polytrauma patients the period will probably be longer. Therefore, the use of helicopters with medical crews will probably not make the period of production loss shorter than the friction period. Hence, it is not to be expected that net yields in production will occur.

In the current analysis it is estimated that, in the baseline scenario, 29 extra patients will have to survive compared to the current situation to make the service cost-effective, if DFL 50,000 is accepted as the limit for cost-effectiveness. Irrespective of the question whether the assumptions on which this scenario is based are correct - which is not certain because of the scarcity of data - the question presents itself whether such favourable effects can actually be attained in practice.

In the estimation of the gain in survival that can be expected to occur, it is important to realise that not all *in-hospital deaths* in seriously wounded accident victims can be prevented.[29] Several papers were published about the gain in survival that can be reached by treatment in trauma centres. Virtually all reported

an improvement in survival.[30-34] Partly because these publications were based on comparisons of mortality before and after introduction of trauma centres, simple conclusions about the percentage improvement in survival are difficult to reach. In the Netherlands the number of persons who died at the scene of a traffic accident or within 30 days of the accident decreased by approximately 48% between 1976 and 1991.[15] If a trauma centre would have been founded in the Netherlands in 1976 an imaginary effectiveness of this centre would have been shown. In such dynamic processes it is very difficult to assess the effectiveness of services and changes in policy in a reliable way.

Also, the maximum gain in survival that can be attained for *extramural mortality* is uncertain.[35,36] In the Netherlands ambulance-nurses are well educated and highly experienced. Because of the ever increasing use of protocols, they also have a large competence. The medical interventions that can be performed at the scene of the accident and during transport are limited, due to the primitive circumstances.[5,8,19] A physician might only be able to set a better indication for performing or omitting certain actions. Comparisons of helicopter services staffed by physicians with helicopter services staffed by paramedics, do not always show better results.[37,38,39,40]

The above illustrates that the factual improvement in survival that can be obtained by a helicopter service with a medical crew is unknown. From the data that are currently available it can not be inferred whether such a service would be cost-effective in the Netherlands. Too many data are missing to make peremptory statements. Especially data on the costs of hospital admissions for seriously wounded trauma patients, the relation between certain forms of medical assistance and the final quality of life of the patient, and data on intramural and extramural preventable mortality are missing. The results presented here leave enough room for doubt to contradict supporters of the service, who say that the cost-effectiveness is so evident that further research is not needed. But adversaries, who say that helicopters with a medical crews are unnecessary in the Dutch situation are not proved right either.

References

1 Muhr G, Kayser M. Mehrfachverletzungen - Rettungssysteme, Bergung und Erstversorgung. Chirurg 1987; 58: 625-30

2 Schmidt U, Frame SB, Nerlich ML, Rowe DW, Enderson BL, Maull KI, Tscherne H. On-scene helicopter transport of patients with multiple injuries - comparison of a German and an American system. J Trauma 1992; 33(4): 548-55

3 Bouillon B, Troidl H, Neugebauer E, Tiling T, Schweins M, Vorweg M. The Cologne prehospital trauma system: results and criteria for its evaluation. Theor Surg 1990; 5: 36-43

4 College voor ziekenhuisvoorzieningen. Advies inzake traumazorg. Rapportnummer 318. Utrecht: College voor Ziekenhuisvoorzieningen, 1991.

5 Goris RJA. De acute zorg voor ernstig gewonde patiënten. Ned Tijdschr Geneeskd 1990; 134(30): 1439-42

6 Onrust JD. Vliegensvlugge dokters. Kampioen 1993; juni: 52-54

7 Tweede Kamer der Staten-Generaal, vergaderjaar 1992-1993. Notitie traumazorg, 23 128, nr. 1. 's-Gravenhage: SDU Uitgeverij Plantijnstraat, 1993

8 Staniforth P. Medical Teams, Flying Squads - the arguments revisited. Injury 1990; 21: 204

9 Champion HR, Sacco WJ, Gainer PS, Patow SM. The effect of medical direction on trauma triage. J Trauma 1988; 28(2): 235-9

10 Goris RJA. Traumatologie in Nederland (brief). Ned Tijdschr Geneeskd 1992; 136(27): 1320

11 Linden AJ van der. Traumatologie in Nederland. Ned Tijdschr Geneeskd 1992; 136(26): 1235-7

12 Moylan JA, Fitzpatrick KT, Beyer AJ3, Georgiade GS. Factors improving survival in multisystem trauma patients. Ann Surg 1988; 207(6): 679-85

13 Patka P, Haarman HJTM, Bakker FC. Traumascoringssystemen intra- en extramuraal. De Ambulance 1992; Oktober: 12-13

14 Stapert JWJL. Medisch zorg voor ernstig gewonden. Is de traumascore bruikbaar in Nederland? Medisch Contact 1991; 22(46): 700-1

15 Centraal bureau voor de statistiek. Hoofdafdeling statistieken van verkeer en vervoer. Statistiek van de verkeersongevallen op de openbare weg 1991. 's-Gravenhage: SDU/uitgeverij, CBS-publicaties, 1992

16 ÖAMTC. Die Christophorus-Notarzthubschrauber des ÖAMTC. Einsatzstatistik Jänner - Juni 1991 & Vergleichzahlen 1990. Wenen: ÖAMTC, 1991

17 Draaisma JMTh. Evaluation of trauma care - with emphasis on hospital trauma care - Proefschrift Katholieke Universiteit Nijmegen, 1987.

18 ADAC. ADAC Rettungshubschrauber-Einsatzstatistik 1992. München: ADAC-Luftrettung GmbH, 1993

19 Wet op de lijkbezorging. Koninklijk besluit 3 mei 1991, Stb 252.

20 Centraal bureau voor de statistiek. Maandstatistiek van de bevolking. December 1992. 's-Gravenhage: SDU/uitgeverij, CBS-publicaties, 1992

21 Berden B, Hughes B, Magazu D, Nunen, Dijkstra M. An overview of the introduction of airborne E.M.S. into the Netherlands. Rotterdam: Rotterdam School of Management, 1987

22 Charro FTH de. Kosten-effectiviteitsanalyse van het nierfunctie-vervangings-programma in Nederland. Delft: Eburon, 1988.

23 Hout BA van, Bonsel GJ, Habbema JDF, Maas PJ van der, Charro FTH de. Heart transplantation in the Netherlands: costs, effects and scenarios. J Health Economics 1993; 12: 73-93

24 Michel BC, Bonsel GJ, Stouthard MEA, McDonnell J, Habbema JDF. Kosten-effectiviteitsanalyse van levertransplantatie op lange termijn; het Groningse levertransplantatieprogramma 1979-1991. Ned Tijdschr Geneeskund 1993; 137(19): 963-9

25 Rutten FFH, Ineveld BM van, Ommen R van, Hout BA van, Huisman R. Kostenberekening bij gezondheidszorgonderzoek. Richtlijnen voor de praktijk. Rijswijk: Stuurgroep Toekomstscenario's Gezondheidszorg Uitgeverij Jan van Arkel, 1993

26 Koopmanschap MA, Rutten FFH. Indirect costs in economic studies. Confronting the confusion. PharmacoEconomics 1993; 4(6): 446-54

27 Koopmanschap MA, Rutten FFH, Ineveld BM van, Roijen L van. The friction cost method for measuring indirect costs of disease. J Health Economics (in press)

28 Centraal Bureau voor de Statistiek. Diagnose-statistiek bedrijfsverenigingen 1990. Omslagleden. Voorburg/Heerlen: Centraal Bureau voor de Statistiek, 1993

29 Draaisma JM, Haan AFJ de, Goris RJA. Preventable trauma deaths in the Netherlands - a prospective multicenter study. J Trauma 1989; 29(11): 1552-7

30 Champion HR, Sacco WJ, Copes WS. Improvement in outcome from trauma center care. Arch Surg 1992; 127: 333-8

31 Shackford SR, Hollingworth-Fridlund P, CooperGF, Eastman AB. The effect of regionalization upon the quality of trauma care as assessed by concurrent audit before and after institution of a trauma system: a preliminary report. J Trauma 1986; 26(9): 812-20

32 West JG, Cales RH, Gazzaniga AB. Impact of regionalization. The Orange County Experience. Arch Surg 1983; 118: 740-4

33 Kane G, Wheeler NC, Cook S, Englehardt R, Pavey B, Green K, Clark O, Cassou J. Impact of the Los Angeles County Trauma system on teh survival of seriously injured patients. J Trauma 1992; 32(5): 576-83

34 Rutledge R, Messick J, Baker C, Rhyne S, Butts J, Meyer A, Ricketts T. Multivariate population-based analysis of the association of county trauma centers with per capita county trauma death rates. J Trauma 1992; 33(1): 29-37

35 Baxt WG, Moody P. The impact of advanced prehospital emergency care on the mortality of severely brain-injured patients. J Trauma 1987; 27(4): 365-69

36 Gilroy D. Deaths (144) from road traffic accidents occurring before arrival at hospital. Injury 1985; 16: 241-2

37 Schiller WR, Knox R, Zinnecker H, Jeevanandam M, Sayre M, Burke J, Young DH. Effect of helicopter transport of trauma victims on survival in an urban trauma center. J Trauma 1988; 28(8): 1127-34

38 Baxt WG, Moody P. The impact of a physician as part of the aeromedical prehospital team in patients with blunt trauma. JAMA 1987; 257(23): 3246-50

39 Hamman BL, Cue JI, Miller FB, O'Brien DA, House T, Polk HC Jr, Richardson JD. Helicopter transport of trauma victims: does a physician make a difference? J Trauma 1991; 31(4): 490-4

40 Snow N, Hull C, Severns J. Physician presence on a helicopter emergency medical service: necessary or desirable? Aviation, Space and Environmental Medicine 1986; 57: 1176-8

THE ROLE OF A CLINICAL DECISION RULE IN SYMPTOMATIC PULMONARY EMBOLISM PATIENTS WITH A NON-HIGH PROBABILITY VENTILATION-PERFUSION SCAN*

Summary

In this chapter the construction of a simple, quantitative clinical decision rule (CDR) for the diagnosis of pulmonary embolism (PE) in patients with a non-high probability ventilation-perfusion scan is illustrated. The CDR was constructed by multivariate stepwise logistic regression analysis on 28 clinical and lung scan items gathered prospectively and independently in consecutive symptomatic patients. Data on 140 consecutive patients admitted to the Academic Medical Centre and the Slotervaart Hospital in Amsterdam who had a non-high probability result of the ventilation perfusion scan and an interpretable pulmonary angiogram were used. The prevalence of a proven pulmonary embolism was 27.1%. A clinical decision rule containing the presence of wheezing, previous deep venous thrombosis, recently developed or worsened cough, body temperature above 37°C and multiple defects on the perfusion scan was constructed. For the clinical decision rule the area under the Receiver Operating Characteristic (ROC) curve was larger than that of the prior probability of pulmonary embolism as assessed by the physician before lung-scanning.(0.76 versus 0.59; p = 0.0097) At the cut-off point with the maximal positive predictive value 2% of the patients scored positive, at the cut-off point with the maximal negative predictive value pulmonary embolism could be excluded in 16% of the patients. The optimal use of the rule appears to be in the exclusion of PE. Prospective validation of this rule is indicated to confirm its clinical utility.

* Based on: Michel BC, Kuijer PMM, McDonnell J, Beek EJR van, Rutten FFH, Büller HR. The derivation of a clinical decision rule in symptomatic pulmonary embolism patients with a non-high probability ventilation-perfusion scan. Submitted.

Introduction

Ventilation-perfusion lung scanning is a reliable and frequently used diagnostic method in patients with clinically suspected pulmonary embolism. Therapeutic decisions can be based on a negative or high probability result. However, in more than 40% of the patients a non-high probability scan result is obtained. To confirm or exclude the presence of pulmonary embolism in these patients, of whom 20 to 30% have the disease, can be very difficult.[1] For this goal multiple diagnostic tests, including serial noninvasive leg testing for venous thrombosis, pulmonary angiography, blood tests for the detection of coagulation activation and clinical assessment, are advocated.[2-4]

Recently, it was shown that the addition of clinical assessment to the interpretation of lung scan results can increase the diagnostic accuracy for pulmonary embolism.[2,5] Since the clinical information in that study was, however, gathered in an unstandardised fashion, the reproducibility of the results may be hampered. In general there is considerable variation among physicians in the way clinical information is utilised in the diagnostic process of pulmonary embolism.[6] Since this variation does not seem to decrease with increased experience of the physician, training will probably not improve the use of this information.[6] Standardising the clinical assessment appears to be the best solution. Several investigators have tried to derive clinical decision rules in order to reach such standardisation.[7-9] These decision rules generally contain a number of clinical items, each with a weight indicating their predictive value for the presence or absence of pulmonary embolism. After the weighted scores of the items have been aggregated in a total score, the clinical decision rule can be used like a laboratory test: if the total score is above a certain level (the cut-off point), the presence of pulmonary embolism is highly likely, and on the other hand if it is below a certain level, the presence of pulmonary embolism is virtually excluded. However, to make these decision rules reliable and easy to use in a clinical setting, they should only include reproducible items that can be obtained in a simple way. None of the available decision rules meets these criteria.

The aim of our prospective study in consecutive symptomatic patients was to derive a simple and easily reproducible clinical decision rule that can improve the current use of diagnostic information in patients with a non-high probability result of the ventilation-perfusion scan.

Patients and methods

PATIENTS

Of 452 consecutive patients with clinically suspected pulmonary embolism who participated in a prospective clinical study in the Academic Medical Centre and the Slotervaart Hospital in Amsterdam, the Netherlands, a clinical registration form was completed by the physician directly in charge of the patient. Only data on patients with a subsequent non-high probability ventilation-perfusion scan and an interpretable pulmonary angiogram were used to construct the clinical decision rule. After the recording of the history, physical examination, chest X-ray and routine laboratory tests, but before lung scanning, the same physician was asked to score the patient's probability of having pulmonary embolism as less than 10%, 10 to 50%, 50 to 90%, or more than 90%.

CLINICAL DATA

The clinical registration form contained questions on demographic characteristics, the present and past medical history and the physical examination. For the derivation of the clinical decision rule this information and data related to qualitative aspects of the perfusion scan (the presence of multiple defects and the presence of segmental defects) were used. Items on which only a small percentage of patients scored positive (e.g. leg paresis, recent plaster cast or childbirth), with a definition that was found to be unclear (e.g. a third or fourth heart sound, heart failure and apprehension/fear) or that had frequent missing data (e.g. length, weight and arterial blood gas analysis), indicating that the item was difficult to obtain in practice, were omitted from the analysis. After evaluation, 28 items were thought to be clearly defined and easily reproducible. These items were used to derive the clinical decision rule.

DERIVATION OF THE CLINICAL DECISION RULE

The clinical decision rule was based on a stepwise multiple logistic regression, with the result of the angiogram as the dependent variable. In such

regression analysis a combination of variables is sought that can optimally predict the presence or absence of, in this case, pulmonary embolism. Each variable is given a weight (the coefficient) to indicate its contribution to the prediction of pulmonary embolism. In the application of the rule each variable is scored by multiplying its coefficient by 1 if the variable is present or by 0 if the variable is absent. These scores and the constant are added to reach a total score. The adding of the constant is necessary because patients who have none of the variables in the analysis present will still have a certain chance of having a pulmonary embolism. Depending on the purpose of the rule a cut-off point is chosen above which the rule is called positive. If the rule is used to *rule out* pulmonary embolism, with a maximum level of certainty that the excluded patients indeed do not have pulmonary embolism (a maximum "negative predictive value"), a low cut-off point is chosen. On the other hand if the rule is used to *rule in* pulmonary embolism, with a maximum level of certainty that patients with a positive result indeed have pulmonary embolism (a maximum "positive predictive value"), a high cut-off point is chosen. Usually, in logistic regression the total scores mentioned above are transformed (by log-it transformation) to values between 0 and 1. To prevent difficult calculations, and to make the clinical decision rule more easily applicable in practice, we transformed the values of the cut-off points instead.

In the derivation phase of a clinical decision rule it is important to assess whether the rule will be useful in clinical practice. This depends not only on the clarity of the definitions and on the reproducibility of the variables, but also on the stability of the rule in sub-sets of the patient population. This, of course, should be formally tested in a prospective validation study. Because such validation requires another clinical study, an attempt was made to assess the stability of the selected combination of variables in sub-sets of the patients in the current study. For this purpose, the weights of the variables in the clinical decision rule were recalculated by multiple logistic regression not on the full dataset but on a subset containing a random sample of 73% of these patients (the "derivation patients"). Sometimes derivation was impossible because too few patients with one of the variables present were included in the subset. For all rules that could be derived, the rule with its new coefficients was applied to the remaining 27% of the patients (the "validation patients") and it was evaluated how many of these patients were excluded or included at the points with the maximal positive and negative predictive value, and whether the exclusion or inclusion was correct. This process of "derivation" and subsequent "validation" was repeated 1000 times (using a

computer simulation model built in GLIM 4 (The Numerical Algorithms Group Ltd, Oxford, United Kingdom)), each time with a new random selection of the sub-population for "derivation" and "validation".

STATISTICS

Univariate significance was tested for the continuous variables by Student's two sample *t*-test for variables with a normal distribution or Mann-Whitney rank-sum test for variables without a normal distribution and by Fisher's exact test for the discrete variables. Significance for the Receiver Operating Characteristic curves was tested by the method described by Hanley and McNeil.[10] All statistical analyses were performed in BMDP statistical software release 1990 (BMDP Statistical Software, Los Angeles, California, United States of America).

Results

PATIENTS

A non-high probability result of the ventilation-perfusion scan was obtained in 186 of the 452 consecutive patients (41.2%). In 40 of these patients, pulmonary angiography could not be performed because of heart failure (8 patients), severe pulmonary hypertension (4 patients), poor clinical condition (11 patients), renal failure (3 patients), thrombocytopenia (3 patients), patient's refusal (7 patients), initial misdiagnosis of the scan (3 patients) or logistic reasons (1 patient). In a further 6 patients the result of the angiogram was non-interpretable. Data concerning the remaining 140 patients were used to construct the rule. The mean age of these patients was 57.7 years (SD 17.0), half were males, 60% were seen initially as outpatients. Patients had had symptoms for a median of 2.5 days before entering the diagnostic process. On the angiogram, 38 patients (27.1%) had evidence of pulmonary embolism.

Table 1 *Prevalence of the studied variables from the medical history, physical examination and lung scan findings in the 140 patients with a non-high probability lung scan and an interpretable angiogram.*

	Variable	PE[a] absent N = 102[b]	PE present n = 38	P-value
Medical history[c]	Sanguineous sputum	6.9%	21.1%	0.03
	Cough, new or recently worsened	26.7%	44.7%	0.06
	COPD[d]	33.3%	18.4%	0.10
	Previous deep venous thrombosis	5.9%	13.2%	0.17
	Surgery within the past 3 months	18.6%	7.9%	0.19
	Days of immobilisation	0.0[e]	0.0[e]	0.25
	Malignancy	27.5%	21.1%	0.52
	Gender (percentage male)	52.9%	47.7%	0.57
	Dyspnoea	19.6%	23.7%	0.64
	Previous pulmonary embolism	8.8%	5.3%	0.73
	Palpitations	16.7%	13.2%	0.80
	Days with symptoms	2.0[e]	3.0[e]	0.81
	Age	57.8[f]	57.6[f]	0.93
	Family history of thrombosis	12.7%	10.5%	1.00
	Collapse	4.9%	5.3%	1.00
	Green sputum	7.8%	7.9%	1.00
Physical examination[c]	Wheezing	24.5%	5.3%	0.01
	Body temperature above 37° Celsius	40.2%	60.5%	0.04
	Pleural rub	15.7%	28.9%	0.09
	Breath frequency, in breaths/min	19.3[f]	21.7[f]	0.11
	Leg paresis	2.9%	7.9%	0.34
	Crepitations	37.3%	42.1%	0.70
	Increased central venous pressure	10.8%	13.2%	0.77
	Quetelet-index	24.1[f]	24.0[f]	0.86
	Heart frequency, in beats/min	93.0[f]	92.5[f]	0.90
	Signs of DVT[g]	9.8%	10.5%	1.00
Perfusion lungscan[c]	Multiple defects on perfusion scan	71.6%	83.3%	0.19
	Segmental defects on perfusion scan	55.9%	61.1%	0.70

Legend to table 1: [a] PE = pulmonary embolism as classified by pulmonary angiography; [b] n = number of patients; [c] percentage of patients with variable present; [d] chronic obstructive pulmonary disease; [e] median; [f] average; [g] deep venous thrombosis

CLINICAL DECISION RULE

Most of the investigated variables showed no significant difference for patients with or without pulmonary embolism, table 1. Only body temperature above 37° Celsius and sanguineous sputum were present significantly more often in patients with pulmonary embolism than in those without ($p < 0.05$). Wheezing was more frequent in those shown not to have pulmonary embolism ($p = 0.01$).

In the multivariate stepwise logistic regression, the absence of wheezing, previous deep venous thrombosis, a new or recently worsened cough, a body temperature above 37° Celsius and multiple defects on the perfusion scan were identified as the optimal combination of predicting variables for the presence or absence of pulmonary embolism, table 2. The five selected items each have a weight in predicting the presence or absence of pulmonary embolism in this combination of variables (expressed in the coefficient). The rule is applied by adding the coefficients for the items present and the constant, returning a total score.

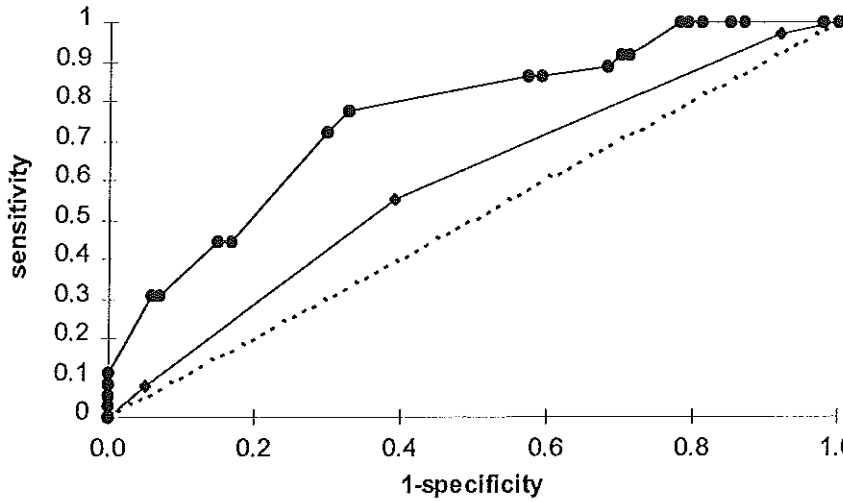
Table 2 *The combination of variables from the medical history, physical examination and lung scan results that has optimal predictive value for the presence or absence of pulmonary embolism in the regression analysis*

Variable	Coefficient	Standard error	P-value
Wheezing	-3.068	1.17	0.0002
Previous deep venous thrombosis	1.886	0.841	0.0210
Cough recently developed or worsened	0.9554	0.445	0.0314
Body temperature above 37°C	0.8847	0.440	0.0413
Multiple defects on perfusion scan	0.9828	0.550	0.0598
Constant	-2.455	0.607	-

Addition of other variables known to be associated with pulmonary embolism, including sanguineous sputum, malignancy, collapse, signs of deep venous

thrombosis and a history of previous pulmonary embolism, or elimination of the weakest predictive variable in the combination, i.e. multiple defects on the perfusion scan, did not improve the predictive value of the rule. Comparison of the performance of the clinical decision rule to that of the prior clinical probability of pulmonary embolism as assessed by the patient's physician at admission, using Receiver Operating Characteristic curve analysis, shows that the clinical decision rule has a better sensitivity and specificity for every cut-off point, figure 1.

Figure 1 *ROC curve of the clinical decision rule (-•-•-•), compared to the ROC curve of the prior clinical probability as assessed at admission by the attending physician (-♦-♦-♦)*



The area under the curve is significantly greater for the CDR than for the clinical assessment (0.76 vs. 0.59; p = 0.0097)

If the rule is used to exclude patients without pulmonary embolism, aiming at a maximum level of certainty that the excluded patients indeed do not have pulmonary embolism, a maximum "negative predictive value" is sought for. A negative predictive value of 100% was reached at the transformed cut-off point of minus 2.512 (0.075 untransformed). Here, the test had a positive predictive value of 31%, a sensitivity of 100% and a specificity of 22%, table 3. This indicates that the rule can be used to exclude pulmonary embolism in 16% of the symptomatic patients with a non-high probability ventilation-perfusion scan.

If the rule is used to select patients with pulmonary embolism, with a maximum level of certainty that patients with a positive result indeed have pulmonary embolism, a maximum "positive predictive value" is sought for. The best positive predictive value that could be reached was 100%, at a transformed cut-off point of 0.405 (0.6 untransformed). At this point a negative predictive value, a sensitivity and a specificity of 75%, 8% and 100% were obtained, respectively. This indicates that the rule is less useful for the selection of patients with pulmonary embolism, since only 2% of the patients with pulmonary embolism can be identified.

Table 3. *Results of the clinical decision rule, based on data from the medical history, physical examination and lungscan findings in 140 patients with a non-high probability lung scan and an interpretable angiogram, compared to the findings on the pulmonary angiogram for the same patients.[a] Numbers are numbers of patients.*

rule used for the exclusion of PE[b]	Pulmonary embolism		
clinical decision rule	present	absent	total
positive	36	79	115
negative	0	22	22
total	**36**	**101**	**137**

rule used for the selection of patients with PE[c]	Pulmonary embolism		
clinical decision rule	present	absent	total
positive	3	0	3
negative	33	101	134
total	**36**	**101**	**137**

[a] *3 patients had a missing value for one of the variables in the rule and had to be excluded*
[b] *cut-off point aimed at maximal negative predictive value of the rule*
[c] *cut-off point aimed at maximal positive predictive value of the rule*

STABILITY

To assess the stability of the combination of variables in separate subpopulations, the coefficients of the variables were estimated in a random selection of 100 "derivation patients" out of the 137 patients who did not have missing data for the variables included in the rule. Subsequently, the rule was "validated" for its usefulness in the exclusion of pulmonary embolism in the remaining 37 patients. Thousand random selections of subpopulations were made, containing in total 100,000 "derivation" patients and 37,000 "validation" patients. The number of patients in whom pulmonary embolism was excluded by the rule in the separate selections of 37 "validation" patients varied between 1 and 17. In total, the rule indicated pulmonary embolism could be excluded in 15.7% (5801 patients) of the "validation" patients. Only 1.3% of these patients (77 patients) did have pulmonary embolism and were, therefore, incorrectly excluded. In conclusion, the chance that limited changes in the patient population in the prospective validation phase will render a clinical decision rule based on this combination of variables invalid seems small.

DISCUSSION

After a surge of interest for technical methods in the diagnosis of pulmonary embolism, there has recently been a revival of attention for the value of information contained in the medical history and physical examination.[2] Both clinical decision rules and neural networks have been advocated in the last few years.[8,9,11]

This study demonstrates that it is possible to construct a simple clinical decision rule that can be used in the diagnostic work-up of patients with suspected pulmonary embolism. If the current clinical decision rule is used to confirm pulmonary embolism in symptomatic patients with a non-high ventilation-perfusion scan, only a few, i.e. approximately 2%, of the patients will be identified. Hence, it seems to be more feasible to use the current decision rule as a cheap and simple test for the exclusion of pulmonary embolism. This appears to be possible in 16% (22 out of 137) of the patients.

The clinical decision rule is easily applicable in the form presented here. Imagine a patient with chronic obstructive pulmonary disease, without a prior

history of deep venous thrombosis, who presents with fever, a recently worsened cough, wheezing and pain related to breathing which instigates the suspicion of pulmonary embolism. A perfusion scan is made which shows multiple defects, and the ventilation-perfusion scan has a non-high probability result. The clinical decision rule is subsequently applied to see whether pulmonary embolism can be ruled out. This patient will have a score of -3.068 (wheezing present) + 0.9554 (recently worsened cough) + 0.8847 (body temperature above 37°C) + 0.9828 (multiple defects on perfusion scan) - 2.455 (constant) = -2.7001, which is below the optimal cut-off point of -2.512 for the exclusion of pulmonary embolism. In this patient the presence of pulmonary embolism is made unlikely by the clinical decision rule. If the patient would have had a previous history of deep venous thrombosis the score would change to -0.8141, which is above the cut-off point. For that patient the clinical decision rule can not help in the exclusion of pulmonary embolism and other diagnostic tests are necessary.

Several aspects of the present investigation warrant comment. The first issue is that the rule was based on data of only 137 patients. In constructing the rule we preferred to base it on data from the patient population in which it will be used, rather than include data on patients with a normal or a high probability ventilation-perfusion scan result. Since the number of variables included in the rule is small, relative to the number of patients, it seems unlikely that over-modelling took place. Furthermore, although the number of patients was limited, they were consecutive and the study was controlled by angiography.

Secondly, some clinical symptoms and signs that are frequently cited in textbooks to be associated with pulmonary embolism were not included in the model. In presenting such criteria, textbooks may be focusing on the total patient population presenting with symptoms of pulmonary embolism, rather than on specific sub-populations, such as those with non-diagnostic lung scan findings. Table 1 shows that many of these criteria do not discriminate between patients with and without pulmonary embolism in an univariate analysis in the patient population studied here. Forcing the well-known criteria into the multivariate regression analysis did not improve the model.

The third issue is that several variables had to be discarded in the process of constructing the rule. Whereas some variables, such as arterial blood gas, heart failure and abnormal heart sounds, may seem feasible beforehand, their definition or reproducibility is cumbersome, causing many missing values or inconsistent

data. The decision rule presented here is based on variables that proved to be collectable in the acute clinical setting of a large teaching hospital, whereas other decision rules frequently contain items that are difficult to reproduce in practice.[8,9]

The last issue is that the clinical decision rule may lose predictive value if the patient population changes. Whereas this may certainly be true for large changes in e.g. the prior probability of pulmonary embolism, the results of the microsimulation in the 1000 random sub-sets of patients point to a reasonable robustness of the combination of variables in sub-populations of the patient population studied.

The current study presents the results of the derivation of a clinical decision rule. The next step should be validation of this rule in a large prospective trial, since earlier studies revealed that the predictive value may decrease significantly upon re-evaluation. Other items for future study could be whether clinical decision rules can add information to the result of other non-invasive tests such as compression ultrasonography or a D-dimer blood test.

We showed that it is possible to derive a simple quantitative clinical decision rule for patients with a non-high probability result of the ventilation-perfusion scan, and that its major application probably lies in ruling out pulmonary embolism.

References

1 Beek EJR van, Tiel-van Buul MMC, Büller HR, Royen EA van, Cate JW ten. The value of lung scintigraphy in the diagnosis of pulmonary embolism. Eur J of Nucl Med 1993; 20: 173-181

2 Stein PD, Hull RD, Saltzman HA, Pineo G. Strategy for diagnosis of patients with suspected acute pulmonary embolism. Chest 1993; 103(5): 1553-9

3 Perrier A, Bounameaux H, Morabia A, Moerloose P de, Slosman D, Unger P, Junod A. Contribution of D-dimer plasma measurement and lower-limb venous ultrasound to the diagnosis of pulmonary embolism: a decision analysis model. Am Heart Journal 1994; 127(3): 624-635

4 Hull RD, Raskob GE, Ginsberg JS, Panju AA, Brill-Edwards P, Coates G, Pineo GF. A noninvasive strategy for the treatment of patients with suspected pulmonary embolism. Arch Intern Med 1994; 154: 289-97

5 Stein PD, Henry JW, Gottschalk A. The addition of clinical assessment to stratification according to prior cardiopulmonary disease further optimizes the interpretation of ventilation/perfusion lung scans in pulmonary embolism. Chest 1993; 104(5): 1472-6

6 Wigton RS, Hoellerich VL, Patil KD. How physicians use clinical information in diagnosing pulmonary embolism: an application of conjoint analysis. Med Decis Making; 6(1): 2-11

7 Petruzzelli S, Palla A, Citi M, Giuntini C. Improvement of screening for pulmonary embolism with a standardized questionnaire. Respiration 1990; 57(5): 329-37

8 Celi A, Palla A, Petruzzelli S, Carrozzi L, Jacobson A, Cella G, Giuntini C, Sasahara AA. Prospective study of a standardized questionnaire to improve clinical estimate of pulmonary embolism. Chest 1989; 95(2): 332-7

9 Hoellerich VL, Wigton RS. Diagnosing pulmonary embolism using clinical findings. Arch Intern Med 1986; 146(9): 1699-704

10 Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983; 148(3): 839-43

11 Patil S, Henry JW, Rubenfire M, Stein PD. Neural network in the clinical diagnosis of acute pulmonary embolism. Chest 1993; 104: 1685-1689

THE COST-EFFECTIVENESS OF DIAGNOSTIC STRATEGIES IN PATIENTS WITH SUSPECTED PULMONARY EMBOLISM*

Summary

The cost-effectiveness of various diagnostic strategies in patients with clinically suspected pulmonary embolism (PE) was analysed using a modelling approach. In 451 consecutive patients with clinically suspected PE data on the performance of diagnostic tests were collected prospectively in two large teaching hospitals in Amsterdam, the Netherlands. The ventilation-perfusion lung scan was used as the primary diagnostic test in all patients. In patients with a non-diagnostic lung scan result the performance of a clinical decision rule, a D-dimer test, and ultrasonography of the leg veins was evaluated with pulmonary angiography as the gold standard. It was estimated that the strategy recommended by a 1992 Dutch consensus meeting costs about DFL 4,400 per patient and that 97.02% of the patients can be expected to survive the first six months after the primary PE. The nation-wide annual costs for the diagnosis and treatment of patients by this strategy were estimated at 163 million Dutch Guilders. Subsequently, the costs and effects of alternative strategies were evaluated in a modelling approach, and compared to those of the consensus strategy. One strategy was selected that produces the best results in terms of survival and leads to considerable savings as compared to the consensus strategy. In this strategy subsequently a ventilation-perfusion scan, a clinical decision rule, a D-dimer test, a pulmonary angiography and leg ultrasonography are performed. Patients with a high probability ventilation-perfusion scan, an abnormal angiography or leg ultrasound test are treated, whereas treatment is withheld in patients with a normal ventilation-perfusion scan, a normal clinical decision rule, a negative D-dimer test, a normal angiography, or a normal leg ultrasound test. This strategy will have to prove its value and usefulness in clinical practice in a subsequent prospective validation phase.

* Based on: Michel BC, Seerden RJ, Beek EJR van, Büller HR, Rutten FFH. The cost-effectiveness of diagnostic strategies in patients with suspected pulmonary embolism. Submitted.

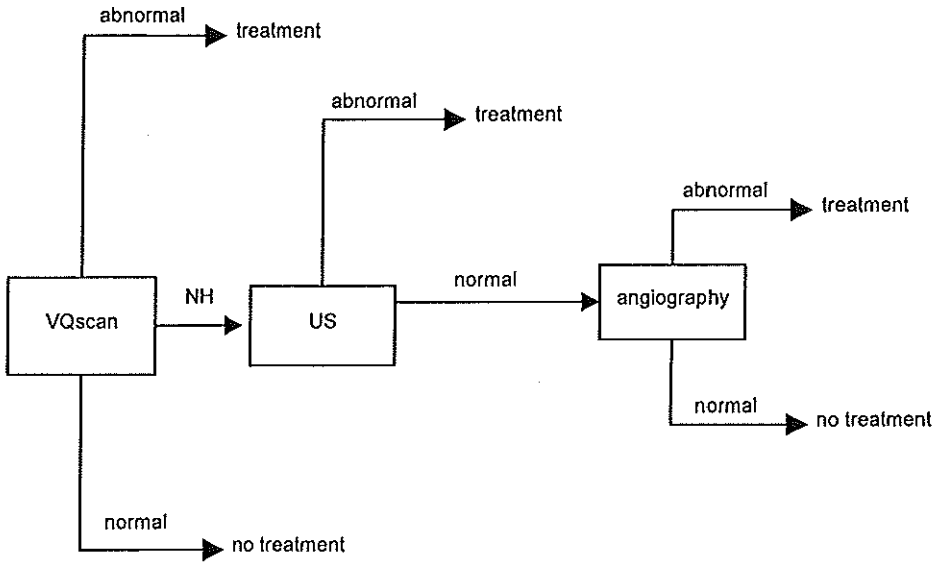
Introduction

Pulmonary embolism (PE) is a serious and potentially life-threatening disease. Together with deep leg vein thrombosis it is regarded as part of the clinical entity venous thromboembolism. The current view-point is that PE is a complication of deep leg vein thrombosis, in which parts of the thrombosis embolise to the systemic circulation and cause complete or partial obstruction of the pulmonary arterial blood flow to the lung. Large emboli may cause immediate death, small emboli may pass unnoticed. If patients surviving the initial PE remain untreated, between 18 and 26 percent may die of recurrent embolism.[1,2] The current treatment policy aims at preventing recurrent embolism, and generally consists of intravenous heparin, followed by three to six months oral anticoagulant therapy. Unfortunately, these preventive treatments may have serious side-effects, most notably fatal haemorrhage.[3,4] Thus, immediate initiation of an appropriate diagnostic strategy is called for, as soon as the diagnosis PE is clinically suspected. Whereas general treatment policies may depend on the incidence of PE in the patient population, for every individual patient suspected of having PE a trade-off between the risks of recurrent embolism and the risks of anticoagulant therapy must be made. For this decision an accurate diagnosis of the presence or absence of PE is warranted.

Diagnosing PE is a challenging problem. Signs and symptoms, like shortness of breath, pain related to breathing, fever, bloody sputum and palpitations, are non-specific and may occur in many other diseases. In most patients the history and physical examination will not be sufficient to diagnose or exclude PE: additional diagnostic tests will be required. For this purpose a perfusion lung scintigraphy is a relatively safe initial diagnostic tool that is widely available.[5] A normal scan effectively rules out PE, but an abnormal scan does not sufficiently prove the diagnosis. In patients with an abnormal perfusion scan a ventilation lung scan usually is the next step in the diagnostic strategy.[6-8] If the combined ventilation-perfusion scan shows at least a segmental perfusion defect with normal ventilation, the scan is designated "high probability", which is generally believed to confirm the disease. But, unfortunately, in approximately 40% to 50% of the patients the result of the ventilation-perfusion scan is inconclusive: a "non-high probability" or non-diagnostic result. In these patients additional diagnostic tests are necessary to prove or exclude PE.

A 1992 consensus meeting of Dutch pulmonologists and internists recommended a policy of ventilation-perfusion scanning, followed by ultrasonography of the deep leg veins in case of an inconclusive ventilation-perfusion scan and a pulmonary angiography in case of a normal leg ultrasound as the most appropriate diagnostic strategy in PE, figure 1.[9] This strategy corresponds with recommendations made by authors from other countries.[10,11]

Figure 1 *The strategy recommended by the Dutch consensus meeting in 1992, which was used as the basis for this study. VQscan = ventilation perfusion scan; NH = non-high probability; US = ultrasound*



However, pulmonary angiography - which is seen as the gold standard for the presence or absence of PE - is not always available, requires expertise, and may lead to serious, sometimes even lethal, complications.[12] Therefore, other strategies, including non-invasive tests such as a D-dimer test (a simple blood test aimed at confirming the existence of degradation products from the thrombus) or a clinical decision rule, are continuously being searched for.[13-18]

The object of the current study is to assess the costs and effects of the Dutch consensus strategy and to compare these with the costs and effects of alternative, possibly less invasive, strategies.

Patients and methods

In a prospective study data on the sensitivity and specificity of various diagnostic tests, the costs of diagnostic tests and treatment, the costs and occurrence of complications, and the occurrence and cause of death were collected. Subsequently a model was built combining these data in order to evaluate and compare the costs and effects of diagnostic strategies in patients with clinically suspected PE.

PATIENTS

Between June 1991 and April 1994 all patients (487) seen with suspected PE in two large teaching hospitals in Amsterdam, the Netherlands, were included in a prospective cohort study. Written informed consent was obtained from all patients and the study was approved by the Institutional Review Boards of both hospitals. All patients were scheduled for lung scintigraphy. If the lung scan was "high probability" anticoagulant treatment was initiated (or continued), whereas treatment was withheld if a normal lung scan was obtained.[8] In patients with a "non-high probability" lung scan result, leg ultrasonography (aimed at confirming the presence of deep leg vein thrombosis), pulmonary angiography and several plasma D-dimer tests were performed. In the present analysis only the most informative D-dimer assay (Diagnostica Stago, Asniere, France) was used.[14]

Additionally, at the time of inclusion into the study, clinical data were recorded by the attending physician for the construction of a clinical decision rule. The rule that was made with these data contains previous deep venous thrombosis, a recently developed or worsened cough, body temperature above 37°C, multiple defects on the perfusion scan and the absence of wheezing as predicting variables. Details on the use and construction of this rule are described in chapter 2.

In all patients, data were recorded on the number of hospital days, the diagnostic tests that were performed and their results, the number of days of heparin treatment, the presence of co-morbidity, the number of days of oral anticoagulant therapy, and the occurrence of bleeding complications and recurrent PE during the initial hospitalisation and the 6 months follow-up period. Finally, the date and probable cause of death were registered.

COSTS

The costs of the diagnostic strategies were made up of two components: the volume of medical activities (average number of hospital days per patient, diagnostic tests, days of heparin treatment, etc.) multiplied by their monetary valuation. Thus, only direct costs made within the health care sector were calculated, whereas costs made by the patients and indirect costs, such as production losses, were not taken into account. All costs were calculated in Dutch Guilders (1 DFL = £ 0.40; 1 DFL = $ 0.64; conversion rates October 1995)

Detailed price calculations were performed in the Academic Medical Centre in Amsterdam, the Netherlands, to estimate unit costs reflecting the real use of resources. These calculations included the costs of manpower, materials, equipment and overheads. The costs of manpower were based on the medium salaries in the gross salary scales for Dutch University Hospitals, and included a 35% increase for social security premiums. Costs of equipment concerned depreciation, interest and costs of maintenance. The first two were calculated on the basis of an economical lifetime of 8 years and an interest of 5%, whereas the yearly costs of maintenance were based on expert opinion and estimated at 8% of the initial expense.

Costs for hospital days were separated from costs for treatment and costs for diagnostic tests and based on the hospital cost accounting system of 1992. They were calculated as the weighted average costs for a hospital day in the relevant wards in the Academic Medical Centre, including costs of salaries for medical specialists and overheads.

For the costs of treatment with heparin on the medical ward only the costs of standard laboratory tests and the heparin pump were included. The other components were already included in the costs for hospital days.

The calculation of the costs for treatment with oral anticoagulants was based on data collected from Dutch Thrombosis Services, which are specialised in outpatient anticoagulant treatment. These costs included two consultations per patient per month and an average of 3 tablets of acenocoumarol 1 milligram daily. As length of treatment with oral anticoagulants the average duration of treatment of the patients in the study, with a maximum of 6 months per patient, was included.

Because at least 73% of the patients in the study had some form of, sometimes serious, co-morbidity the length of stay (LOS) in hospital for the diagnosis and treatment of PE had to be separated from the LOS for the co-morbidity. In determining the average LOS for PE, only hospital days that were necessary for the diagnosis of PE and treatment with intravenous heparin were included. If the real LOS was less than three days longer than necessary for diagnosis of PE and treatment with intravenous heparin, it was assumed that these days were necessary for practical purposes related to the pulmonary embolism, and therefore all these days were ascribed to the PE. If the real LOS was more than 3 days longer, only days necessary for diagnosis of PE and treatment with intravenous heparin were ascribed to the PE, and it was assumed that the other days were related to the patient's co-morbidity.

The average costs of a bleeding complication were estimated by multiplying the length of stay in hospital for each complication with the average cost of a hospital day and adding charges for standard procedures, depending on the site of the bleeding. For cerebrovascular bleeding average real costs were used, based on a previous study (unpublished data). The average costs of diagnosing and treating recurrent clinical PE were estimated by multiplying the length of stay in hospital with the average costs of a hospital day and adding the average costs of diagnostic and therapeutic interventions of a primary PE. The average costs of a hospital day for complications were estimated to be identical to the average costs of a hospital day for a primary PE.

EFFECTS

Effects were calculated as the 6 month survival probability. Complications of treatment, mainly bleedings, and re-occurrence of PE were based on their occurrence in the study patients. Minor complications that would not lead to additional diagnostic or therapeutic procedures or hospital days, e.g. a small haematoma, were not taken into account. Mortality of the angiography was based on literature and estimated at 0.2%.[12,19]

THE MODEL

The model was based on the observed sensitivity and specificity of the separate diagnostic tests in the patients who participated in the study, with the

angiography as the gold standard. Furthermore, data on the incidence of PE in the patient population and the occurrence of PE-related deaths and major complications were included. The model was a decision tree constructed in a spreadsheet programme (Quattro Pro for Windows, version 5.0, Borland International, Inc, 1993) to enable full flexibility. Bayes' theorem was used to calculate the joined sensitivity and specificity for combinations of tests, assuming independence of the various diagnostic tests. In the model diagnostic strategies were applied to 100 imaginary patients. For each diagnostic strategy all possible combinations of test results were defined, and the number of patients with each combination of test results and the probability of their having PE were calculated. Then the costs for tests, treatment and complications, and the survival of patients, based on the mortality of treatment (if given) and the mortality of recurrent PE, were calculated for each combination. And finally the average costs and survival for the diagnostic strategy were calculated by summing the sub-totals for all combinations and dividing the thus obtained results by 100. This process was repeated to calculate the costs and effects of all possible diagnostic strategies.

Both the clinical decision rule and the D-dimer test had variable sensitivity and specificity depending on the cut-off points that were chosen. For the D-dimer nine cut-off points, at every 100 mg/l between 100 mg/l and 900 mg/l, were included in the analysis. For the clinical decision rule all sixteen cut-off points at which the sensitivity and/or specificity of the rule changed were included in the analysis.

Strategies in which the clinical decision rule or the D-dimer test were scheduled to be performed in case the angiography was unavailable or inconclusive were not included in the final analysis, because the sensitivity and specificity of the tests performed in this specific sub-population of patients were thought to be unreliable. All other scenarios were arranged according to the expected gain in survival and the additional costs relative to the Dutch 1992 consensus strategy.[9] Scenarios which were dominated by scenarios involving less costs with an equal or better survival probability, were omitted from the final evaluation. For each of the remaining "dominant" strategies a gain in survival can be attained at certain costs in comparison with the previous strategy. The ratio between the additional costs and the additional effects of a strategy relative to the previous strategy was called the incremental cost-effectiveness ratio (incremental CE-ratio).

The variables used in the calculations of the costs and effects of the various strategies are surrounded by uncertainty margins. The effects of these uncertainties in the final outcome were evaluated in a sensitivity analysis.

Results

A total of 487 consecutive patients were included in the clinical study. For 36 patients no data were available for the cost-effectiveness analysis. Table 1 presents the baseline clinical and sociodemographic characteristics of the remaining 451 patients.

Table 1 *The baseline clinical and sociodemographic characteristics of the 451 patients included in the cost-effectiveness analysis*

Age; average (SD)	56.4 (17.6)[a]
Male/female	193 / 258
Co-morbidity (%)	
-Chronic Obstructive Pulmonary Disease	93 (20.6%)
-Malignancy	103 (22.8%)
Patients referred by the emergency department (%)	266 (59.0%)
Risk factors for pulmonary embolism (%)	
-Recent surgery	95 (21.1%)
-History of venous thrombo-embolism	33 (7.3%)[b]
Signs of deep leg vein thrombosis (%)	39 (8.7%)[c]
Days of symptoms; average (SD)	7.2 (14.7)[d]

[a] *446 patients;* [b] *450 patients;* [c] *449 patients;* [d] *443 patients*

On average two hospital days were necessary for diagnosing or excluding PE in these patients. Four-hundred-and-forty-four of these patients had a ventilation-perfusion scan, 126 (28%) with a normal result and 132 (30%) with a high probability result. In the remaining 186 (42%) patients the result of the ventilation-perfusion scan was inconclusive.

In 40 of the 186 patients with a "non-high probability" result of the ventilation-perfusion scan angiography could not be performed for various technical reasons, and in 6 additional patients the angiogram could not be interpreted. In the

remaining 140 patients 38 patients had an abnormal angiogram, hence the prevalence of PE in patients with a "non-high probability" lung scan was 27%.

Table 2 gives an overview of the sensitivity and specificity of the various tests, with the pulmonary angiography as the gold standard.

Table 2 Sensitivity, specificity and 95% confidence interval (CI) for the ventilation-perfusion scan, the leg ultrasound and several cut-off points ("cut-off") for the D-dimer test and the clinical decision rule

Test	cut-off	sensitivity (CI)	specificity (CI)
Ventilation-perfusion scan[a]	-	0.68 (0.61-0.75)	0.44 (0.38-0.50)
Deep leg vein ultrasonography[b]	-	0.29 (0.12-0.45)	0.96 (0.90-0.99)
D-dimer[c]	300	1.00 (0.90-1.00)	0.12 (0.03-0.20)
	500	0.89 (0.72-0.98)	0.29 (0.18-0.41)
	700	0.75 (0.57-0.93)	0.44 (0.32-0.57)
Clinical Decision Rule[d]	0.075	1.00 (0.91-1.00)	0.22 (0.13-0.30)
	0.192	0.78 (0.63-0.93)	0.67 (0.58-0.77)
	0.358	0.44 (0.27-0.62)	0.83 (0.75-0.91)
	0.592	0.11 (0.00-0.23)	1.00 (0.97-1.00)

[a] *In 444 patients with a ventilation-perfusion scan. The prevalence of PE in high probability (88%) and negative scans (4%) is based on data from literature.[8] It is assumed that the prevalence of PE in the 46 patients with a non-high probability scan, but without an interpretable angiography, is identical to the prevalence in the 140 patients with an interpretable angiography (27%).*

[b] *In 131 patients with a non-high probability ventilation-perfusion scan, a leg ultrasound and an interpretable angiography*

[c] *In 96 patients with a non-high probability ventilation-perfusion scan, a Stago D-dimer test and an interpretable angiography*

[d] *In 137 patients with a non-high probability ventilation-perfusion scan, an interpretable angiography and sufficient clinical data to complete the clinical decision rule*

Treatment with oral anticoagulants was given for 3.7 months on average. The average duration of treatment with heparin was 5.4 days. Two of the 38 patients who were treated with anticoagulant therapy after an abnormal angiogram subsequently had a lethal recurrent PE within 6 months (5.3%). The number of clinical recurrences, warranting hospital admission and additional diagnosis and thereby causing additional costs, was 2.6 times the number of lethal recurrences.

One patient died due to a fatal haemorrhage during heparin treatment, while heparin was used for a total of 1806 patient days. Nine patients had a major bleeding episode during a total of 709 patient months of oral anticoagulant treatment, one of these patients died. Table 3 gives an overview of the costs per unit, which were used in the calculations.

Table 3 *Costs (in Dutch Guilders) per unit as included in the model.*

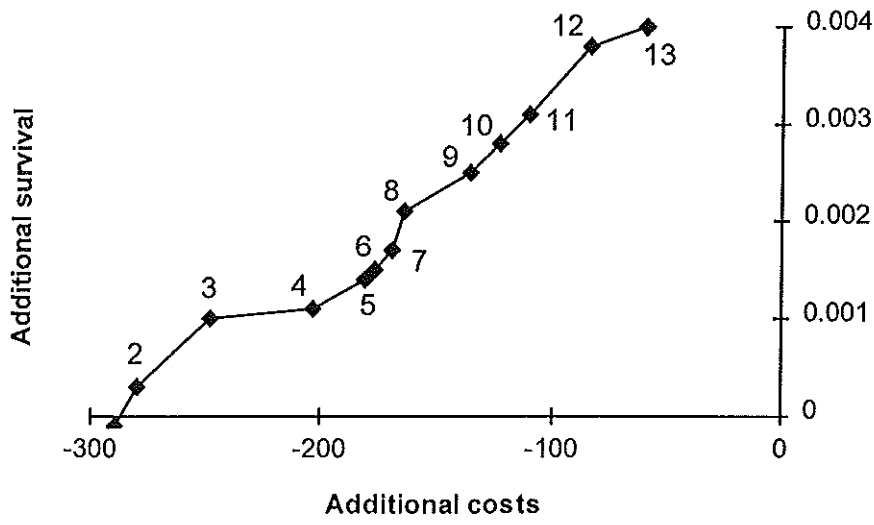
Item		Costs
Hospital day	DFL	568.00
Treatment with heparin: 1st day	DFL	26.83
following days	DFL	13.23
Treatment with acenocoumarol per month	DFL	42.33
D-dimer test	DFL	43.00
Leg ultrasound	DFL	68.00
Clinical decision rule	DFL	40.00
Pulmonary angiography	DFL	765.00
Perfusionscan	DFL	313.00
Ventilation-perfusionscan	DFL	549.00
Bleeding complication	DFL	6,601.56
Recurrent pulmonary embolism	DFL	4,014.00

The model showed that diagnosing (or excluding) PE with the consensus strategy costs approximately DFL 4,398.23 per patient and that 97.02% of the patients can be expected to survive the first 6 months after the primary PE. These results were compared with the costs and effects of 11,605 alternative scenarios. Subsequently 11,563 scenarios that were dominated by better ones were omitted. In 12 of the remaining 42 "dominant" strategies less investment than in the consensus strategy yielded more survival.

Table 4 gives an overview of the costs and effects of these 12 strategies. For these 12 strategies the incremental CE-ratio relative to the previous strategy is presented in a separate column. In this overview a large incremental CE-ratio means that large investments will have to be made to gain additional survivors relative to the previous strategy. It will not be efficient to choose a strategy with a

high incremental CE-ratio relative to the previous strategy (like strategy 4). Neither will it be efficient to use a strategy that is followed by a low incremental CE-ratio (like strategy 7). Figure 2 shows the non-decreasing function of survival relative to costs for the dominant strategies relative to the consensus strategy. The strategies on top of the small notches of the line are the most appealing strategies, like strategies 3, 8 and 12. These strategies combine relatively low costs with a high survival in comparison with the neighbouring strategies.

Figure 2 *The non-decreasing function of survival relative to costs for the 12 dominant strategies, compared to the consensus strategy. The numbers near the line indicate the number that was given to the strategy in table 4.*



The last column of table 4 shows the nation-wide costs of diagnosing and treating 37,000 patients annually for each of the diagnostic strategies.[20] The nation-wide costs of diagnosing and treating PE range from 75 (for a strategy in which no tests are performed and no treatment is provided) to 163 million Dutch Guilders per year.

Table 4 *Average costs and effects per patient and nation-wide costs of the 12 "dominant" strategies in patients with a non-high probability ventilation-perfusion scan, which yield a better survival than the consensus strategy. us = ultrasound; angio = angiogram; ddimer400 = D-dimer test with cut-off point 400; cdr0.192 = clinical decision rule with cut-off point 0.192; (-) = if result is negative; (+) = if result is positive; (np.in) = if not possible or inconclusive; pp = per patient.*

	Strategy	costs pp	6 months survival	additional costs	additional survival	incremental CE-ratio	national costs
1	Reference: us(-)angio(np.in)treatment	4,398	0.9702				162,734,510
2	cdr0.192(+)ddimer400(+)angio(np.in)us	4,118	0.9705	-280	0.0003	23,925	152,374,880
3	cdr0.192(+)ddimer300(+)angio(np.in)us	4,150	0.9712	-248	0.0010	45,243	153,546,670
4	cdr0.175(+)ddimer400(+)angio(np.in)us	4,195	0.9713	-204	0.0011	446,200	155,197,610
5	cdr0.125(+)ddimer500(+)angio(np.in)us	4,217	0.9716	-181	0.0014	75,900	156,040,100
6	cdr0.092(+)ddimer500(+)angio(np.in)us	4,221	0.9717	-177	0.0015	37,200	156,177,740
7	cdr0.175(+)ddimer100(+)angio(np.in)us	4,229	0.9719	-169	0.0017	39,200	156,467,820
8	ddimer600(+)cdr0.075(+)angio(np.in)us	4,234	0.9723	-164	0.0021	13,125	156,662,070
9	ddimer500(+)cdr0.075(+)angio(np.in)us	4,262	0.9727	-136	0.0025	69,775	157,694,740
10	ddimer400(+)cdr0.025(+)angio(np.in)us	4,276	0.9730	-123	0.0028	45,100	158,195,350
11	ddimer400(+)cdr0.075(+)angio(np.in)us	4,288	0.9733	-110	0.0031	42,633	158,668,580
12	cdr0.042(+)ddimer300(+)angio(np.in)us	4,314	0.9740	-84	0.0038	37,214	159,632,430
13	cdr0.075(+)ddimer300(+)angio(np.in)us	4,339	0.9742	-59	0.0040	124,000	160,550,030

Table 5 provides an overview of the results of a sensitivity analysis in relation to the strategy with the highest survival (strategy 13). This shows that the number of hospital days has a major influence on the costs. The percentage of treated patients dying of recurrent PE has an important influence on the estimated survival. If this percentage increases from 5.3%, in the baseline scenario, to 6.3%, survival decreases by 0.4%. The percentage of untreated patients dying of recurrent PE turned out to have a smaller impact on survival.

Table 5 *Results of the sensitivity analysis for some of the main variables in the model. (BL:) = value of item in baseline scenario*

	costs in DFL	6 months survival
Baseline scenario	4,339	97.42%
Lethality untreated PE 26% (BL: 18%)	4,349	97.33%
Lethality untreated PE 38%	4,364	97.18%
Lethality recurrent PE treated 6.4% (BL: 5.3%)	4,377	97.05%
Costs bleeding complications DFL 5,600 (BL: DFL 6,602)	4,321	97.42%
Costs bleeding complications DFL 7,600	4,358	97.42%
Costs recurrent PE DFL 3,000 (BL: DFL 4,014)	4,285	97.42%
Costs recurrent PE DFL 5,000	4,392	97.42%
Three hospital days necessary for diagnosis (BL: 2)	4,916	97.42%
1 bleeding complication / 100 patient years of treatment (BL: 1.7)	4,289	97.51%
4 bleeding complications / 100 patient years	4,507	97.14%
7 bleeding complications / 100 patient years	4,724	96.78%

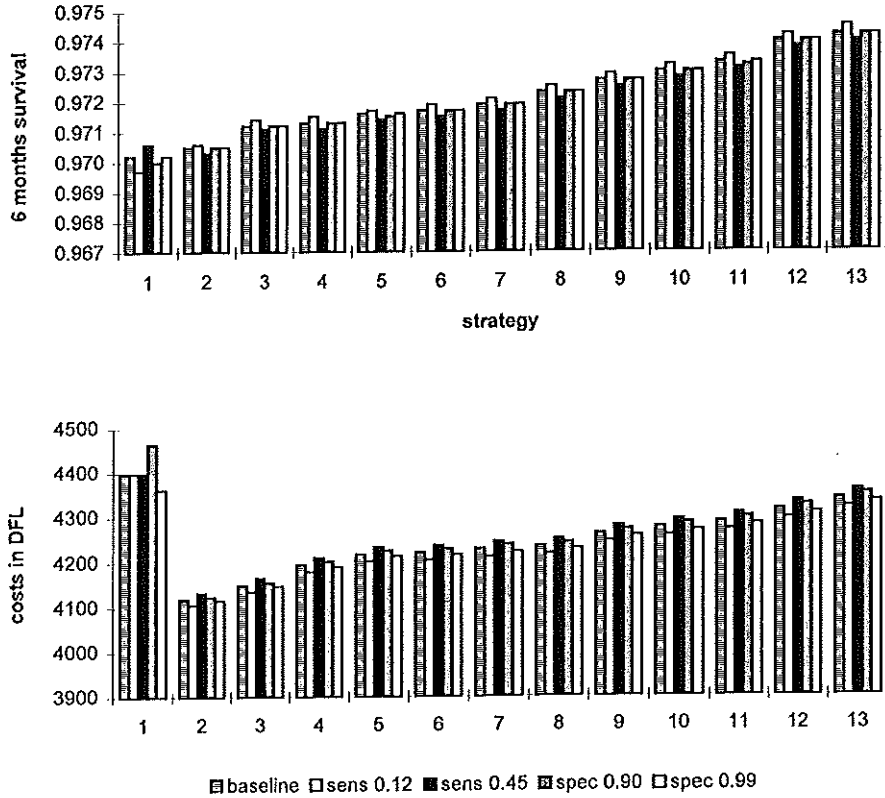
Another major influence on survival is seen in the number of bleedings in 100 patient years of treatment with oral anticoagulants. In this paper the definition of bleeding (any bleeding likely to cause extra costs) was broader than the definition "major bleeding" in a large Dutch study into the complications of anticoagulant treatment.[21] The *total* number of bleeding episodes (15.3 per 100 patient years) in our study was smaller than in patients receiving oral anticoagulants because of non-arterial disease in the above mentioned study (21.1 per 100 patient years), but much larger than the number of *major* bleedings in that study (3.9 per 100 patient

years). The number of *fatal* bleedings was in the same order of magnitude (1.7 per 100 patient years in our study vs. 1.2). Even though table 5 shows that the frequency of bleeding has a major influence both on costs and survival, sensitivity analysis shows that the *ordering* of the 12 strategies will remain unchanged if an increase in fatal bleedings to 7 per 100 patient years occurs in combination with an increase in (both fatal and non-fatal) bleedings to 63 per 100 patient years, or if a decrease in fatal bleedings to 1 per 100 patient years occurs in combination with a decrease in total bleedings to 9 per 100 patient years.

The effects of the confidence intervals surrounding the sensitivity and specificity on the costs and effects for the dominant strategies are presented in figure 3. This figure shows that for the leg ultrasound changes in sensitivity and specificity within the confidence intervals had little influence on the ordering of the 12 dominant strategies.

Here only the influence of the confidence intervals surrounding the sensitivity and specificity of the leg ultrasound are presented. For the other diagnostic tests sensitivity analysis was thought to be less opportune. The ventilation-perfusion scan was used in all patients as the first test, therefore a change in sensitivity or specificity of this test would not influence the ordering of the strategies. For the D-dimer and clinical decision rule many values for sensitivity and specificity were already included in the baseline analysis. And the angiography was used as the gold standard with perfect test results.

Figure 3 *Results of the sensitivity analysis for changes in the sensitivity (sens) and specificity (spec) within the confidence intervals for the leg ultrasound*



DISCUSSION

At present, the evaluation of costs and effects plays a major role in the decision about the distribution of health care services. Whereas the major focus is on the cost-effectiveness of therapeutic technologies, large sums of money are spent on the diagnosis of diseases with a high incidence in the general population. The current analysis shows that under the Dutch consensus strategy approximately DFL 163 million is spent on the diagnosis and therapy of PE annually, amounting to 0.4% of the total costs of health care in the Netherlands in 1988. This is considerably more than the total costs spent on the diagnosis and treatment of

prostate cancer (110 million) or stomach cancer (99 million) in 1988 in the Netherlands.[22]

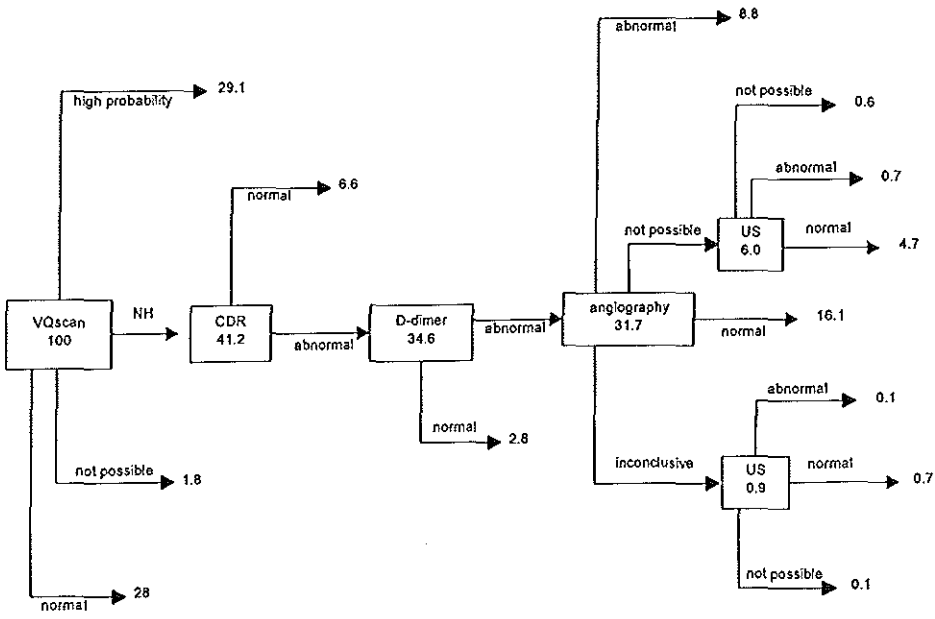
Here, a modelling approach towards the analysis of the cost-effectiveness of diagnostic strategies in pulmonary embolism was chosen to select the most favourable diagnostic strategy from a large variety of possible strategies. Because the selected 12 strategies all show both savings and longer survival compared to the baseline strategy, the incremental cost-effectiveness when considering one strategy after another helps to select the most efficient strategy. More specifically, a strategy is relatively efficient when the incremental cost-effectiveness relative to the previous strategy is low and the incremental cost-effectiveness of proceeding to the next strategy is high.

In order to place the current results in perspective we made a rough calculation of the costs per QALY for each strategy. Survival was converted to QALYs by calculating the average life expectancy for the patients in the study, according to their age and sex with data from the Dutch Central Bureau for Statistics. Because of considerable co-morbidity in the patient population we reduced the life expectancy by an arbitrary 25%. The average quality of life of patients was set at the descriptive scores of the EuroQol questionnaire three months after the PE.[23] These scores were transformed to a normalised 0 to 1 scale using median values of the Dutch general public.[24] The correction factor for quality of life of 0.7 that was thus obtained, was assumed to remain constant over the rest of the patients' life. After discounting by 5%, an average life expectancy of 9.2 QALYs was estimated. This estimation was subsequently used to convert the incremental CE-ratios into incremental costs per QALY. Under these assumptions the maximum incremental CE-ratio equals DFL 48,365 per QALY (strategy 3 to strategy 4). This does not surpass the ratio of DFL 55,000 per QALY, that was deemed acceptable when deciding to introduce liver transplantation and heart transplantation in the Netherlands. However, it should be noted that for less path-breaking technologies the threshold may be lower than DFL 55,000 per QALY. If an incremental CE-ratio of less than DFL 55,000 per QALY is accepted as cost-effective, the strategy with the maximum survival (strategy 13) will be the preferred choice. (Table 4) The CE-ratio of this strategy relative to the consensus strategy, can be estimated at minus DFL 1,600 per QALY.

In strategy 13 subsequently a ventilation-perfusion scan, a clinical decision rule, a D-dimer test, a pulmonary angiography and leg ultrasonography are

performed. Patients with a high probability ventilation-perfusion scan, an abnormal angiography or an abnormal leg ultrasound test are treated with anticoagulants, whereas treatment is withheld in patients with a normal ventilation-perfusion scan, a normal clinical decision rule, a negative D-dimer test, a normal angiography, or a normal leg ultrasound test. Figure 4 gives an overview of the numbers of patients with normal or abnormal test results, according to the model, for strategy 13. This strategy combines the various available diagnostic methods in a clinically useful way by performing simple non-invasive tests first (such as the ventilation-perfusion scan, the clinical decision rule and the D-dimer test). Hence, pulmonary angiography, a method which requires cardiac catheterisation, is limited to patients with non-diagnostic non-invasive test results.

Figure 4 *The diagnostic strategy that produces the best results in terms of survival in combination with considerable savings relative to the consensus strategy. VQ-scan = ventilation-perfusion scan, CDR = clinical decision rule, US = ultrasonography, angio = angiography, + = abnormal test result, - = normal test result, NH = non-high probability result. The numbers are the numbers of patients with the specified combination of test results, predicted by the model for an imaginary 100 patients.*



In estimating costs of diagnosis and treatment of PE, co-morbidity plays a major role. In the patient population analysed 73% of the patients had some form of concurrent disease, implying that all costs for PE had to be separated from the costs made for the concurrent disease. In this analysis, the arbitrary decision was made to attribute all hospital days to PE in which diagnostic or therapeutic procedures for PE were performed. These costs are overestimated if patients had to stay in hospital anyway because of their co-morbidity.

Another major assumption in the current analysis was that of the independency of the tests. We felt that this assumption was justified because the tests included in the analysis all had their impact at different points of the thrombosis forming process. The effect of this assumption and of missing values for the separate tests was analysed by applying strategy 13 to the original database. No important differences were observed. The model predicted that 31.7% of the patients would need pulmonary angiography for an accurate diagnosis, whereas application of the strategy to the original database showed that 32.1% of the patients would have to undergo this test. With the model it was predicted that 6.6% and 2.8% of the patients would not have to be treated because of a negative clinical decision rule or a normal D-dimer test. In reality this was the case in 7.5% and 1.5% of the patients, respectively.

Despite the fact that large sums are spent on the diagnosis and treatment of PE, mainly because of its high incidence, to our knowledge only one previous study on the costs and effects of PE was performed. That study by Oudkerk et al was not based on (prospective) patient data and did not include detailed cost calculations nor quality of life data.[25] The study resulted in a recommendation to apply the consensus strategy as the most cost-effective strategy. The current analysis shows that the addition of non invasive tests may improve considerably on this strategy.

The next step will be to validate the most favourable strategy in a prospective study in a clinical setting. Especially the sensitivity and specificity of the clinical decision rule requires further validation.

References

1 Barritt DW, Jordan SC. Anticoagulant drugs in the treatment of pulmonary embolism, a controlled trial. Lancet 1960; 1309-12

2 Coon WW, Willis PW, Keller JB. Venous thromboembolism and other venous disease in the Tecumseh community health study. Circulation 1973; 48: 839 - 846

3 Levine MN, Raskob G, Hirsh J. Hemorrhagic complications of long-term anticoagulant therapy. Chest 1989; 95 suppl: 26S - 36S

4 Isselbacher KJ, Braunwald E, Wilson JD, Martin JB, Fauci AS, Kasper DL. Harrison's principles of Internal Medicine. New York: Mc Graw-Hill Inc, 1994.

5 Beek EJR van, Tiel-van Buul MMC, Buller HR, Royen EA van, Cate JW ten. The value of lung scintigraphy in the diagnosis of pulmonary embolism. European Journal of Nuclear Medicine 1993; 20: 173-181

6 Gottschalk A, Juni JE, Sostman HD, Coleman RE, Thrall J, McKusick KA, Froelich JW, Alavi A. Ventilation-perfusion scintigraphy in the PIOPED study. Part I. Data collection and tabulation. Journal of Nuclear Medicine 1993; 34: 1109-1118

7 Gottschalk A, Juni JE, Sostman HD, Coleman RE, Thrall J, McKusick KA, Froelich JW, Alavi A. Ventilation-perfusion scintigraphy in the PIOPED study. Part II. Evaluation of the scintigraphic criteria and interpretations. Journal of Nuclear Medicine 1993; 34: 1119-26

8 The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). Journal of the American Medical Association 1990; 263: 2753-9

9 Briët E, van Beek EJR, Oudkerk M. Consensus on the diagnosis of pulmonary embolism [in Dutch]. Nederlands Tijdschrift voor Geneeskunde 1993; 137: 255-9

10 Hull RD, Raskob GE, Hirsh J. The diagnosis of clinically suspected pulmonary embolism. Practical Approaches. Chest 1986; 89(5): 417S-425S

11 Stein PD, Hull RD, Saltzman HA, Pineo G. Strategy for diagnosis of patients with suspected acute pulmonary embolism. Chest 1993; 103(5): 1553-9'

12 Stein PD, Athanasoulis C, Alavi A, Greenspan RH, Hales CA, Saltzman HA, Vreim CE, Terrin ML, Weg JG. Complications and validity of pulmonary angiography in acute pulmonary embolism. Circulation 1992; 85: 462-8

13 Bounameaux H, Cirafici P, de Moerloose P, Schneider PA, Slosman D, Reber G, Unger PF. Measurement of D-dimer in plasma as diagnostic aid in suspected pulmonary embolism. Lancet 1991; 337: 196-200

14 Beek EJR van, Ende B van den, Berckmans RJ, Heide YT van der, Brandjes DPM, Sturk A, Cate JW ten. A comparative analysis of D-dimer assays in patients with clinically suspected pulmonary embolism. Thrombosis and Haemostasis 1993; 70: 408-413

15 Harrison KA, Haire WD, Pappas AA, Purnell GL, Palmer S, Holdeman KP, Fink LM, Dalrymple GV. Plasma D-dimer: a useful tool for evaluating suspected pulmonary embolus. Journal of Nuclear Medicine 1993; 34: 896-8

16 Perrier A, Bounameaux H, Morabia A, Moerloose P de, Slosman D, Unger P-F, Junod A. Contribution of D-dimer plasma measurement and lower-limb venous ultrasound to the diagnosis of pulmonaruy embolism: a decision analysis model. American Heart Journal 1994; 127: 624-35

17 Hoellerich VL, Wigton RS. Diagnosing pulmonary embolism using clinical findings. Archives of Internal Medicine 1986; 146: 1699-704

18 Celi A, Palla A, Petruzzelli S, Carrozzi L, Jacobson A, Cella G, Giuntini C, Sasahara AA. Prospective study of a standardized questionnaire to improve clinical estimate of pulmonary embolism. Chest 1989; 95: 332-7

19 Mills SR, Jackson DC, Older RA, et al. The incidence, etiologies, and avoidance of complications of pulmonary angiography in a large series. Radiology 1980; 136: 295-299

20 Beek EJR van, Büller HR, Everdingen JJE van, Zwijnenburg A, Cate JW ten. Diagnostic management of suspected pulmonary embolism in the Netherlands: results from a questionnaire [abstract]. Thrombosis and Haemostasis 1991; 65: 1171

21 Meer FJM van der, Rosendaal FR, Vandenbroucke JP, Briet E. Bleeding complications in oral anticoagulant therapy. An analysis of risk factors. Archives of Internal Medicine 1993; 153; 1557-1562

22 Koopmanschap MA, Roijen L van, Bonneux L, Barendregt JJ. Current and future costs of cancer. European Journal of Cancer 1994; 30A(1): 60-65

23 The EuroQol Group. EuroQol - A new facility for the measurement of health-related quality of life. Health Policy 1990; 16: 199-208

24 Hout BA van, McDonnell J. Estimating a parametric relation between health description and health valuation using EuroQol instrument. In: Björk S (ed). Discussion paper no 1. EuroQol Conference Proceedings Lund, October 1991. IHE Working paper 1992; 2.

25 Oudkerk M, Beek EJR van, Putten WLJ vab, Büller HR. Cost-effectiveness analysis of various strategies in the diagnostic management of pulmonary embolism. Archives of Internal Medicine 1993; 153: 947-54

ASSESSING THE BENEFITS OF TRANSPLANT SERVICES*

Summary

This chapter presents some methods for the assessment of transplant benefits, including modelling techniques. An independent assessment of the benefits of medical technology is especially important for the evaluation of the balance between the costs and benefits involved. To enable comparison with other health-care facilities, benefits are defined in terms of a combination of life-years gained and quality of life. The number of life-years gained can be calculated by comparing the survival expected with and without transplantation. Survival with transplantation is estimated on the basis of observed survival, acknowledging that the probability of survival may have changed over time, owing to changes in therapy and selection criteria. To estimate survival without transplantation, several techniques are available. Prognostic models, correcting for stage of disease, are often used. Pitfalls in the use of these models are discussed. The number of life-years gained can be corrected for quality of life, by weighing survival with and without transplantation with an index representing quality of life. A method for the calculation of such an index is given. A practical application of the methods described here will be presented in chapter 4B.

Introduction

Nowadays transplant services are provided all over the world for patients with end-stage heart, liver or renal failure. Ideally, new medical technologies develop from an experimental to a non-experimental status in steps that are marked by the results of phase II and III clinical trials. While phase II trials concentrate on activity and safety and try to identify the best treatment protocol in small numbers of

* Based on: Michel BC, Hout BA van, Bonsel GJ. Assessing the benefits of transplant services. Baillière's Clinical Gastroenterology 1994; 8(3): 411-423. Copyright © 1994, by Baillière Tindall.

patients,[1] phase III trials assess the overall benefits of the technology following a fixed protocol and are organised in such a way that statistical evaluation is possible.

Heart, liver and renal transplantation are no longer considered experimental, although their development towards this status has not been characterised by the above mentioned steps. However, some parallels can be observed. After the first, disappointing, results in the 1960s, both surgeons and physicians gained experience in the selection and investigation of patients, the timing of surgery in the course of disease, surgical and anaesthetical techniques, postoperative care and the detection and treatment of rejection. New and better immunosuppressive agents were sought and found,[2] and one might call this stage - which lasted for about 20 years - the equivalent of a phase II trial. A similar equivalent of the randomised controlled phase III trial cannot be found. When in the early 1980s the first favourable post-transplant survival curves were published,[3,4] transplantation rapidly became an established form of therapy. While the number of transplant centres increased steadily, no formal experimental comparison was ever carried out. The obvious reason for this is that, because transplant services showed more and more favourable effects, randomised controlled trials comparing transplantation with alternative strategies rapidly became unethical.[5-7] So, assessment of the benefits has to be based on other types of information.

If safety and efficacy are proven, and if effects are favourable in comparison to those of alternative strategies, what is the relevance of a precise assessment of benefits? The answer is costs. New medical technologies often cause health-care budgets to increase and in most Western countries it is common practice to question whether the additional benefits are worth the additional costs. The concept of "value for money" has been introduced in health care, most notably in the market for pharmaceuticals,[8] but also for other health-care interventions.[9,10] Economic evaluation is acknowledged as the appropriate technique to assess the balance between additional costs and additional benefits. In most circumstances the outcome of such studies is summarised in terms of a so-called cost-effectiveness ratio: additional costs in the numerator and additional benefits (in terms of improvement in health status) in the denominator. However, the demand for cost-effectiveness information has various implications. First, to be able to calculate such a ratio, the knowledge that a therapy has positive effects is not enough; rather, a precise estimate of the effects should be available. Second, a common summary measure of benefits is asked for, that is relevant across all kind

of health-care facilities. As a consequence, it is now common practice to try to express the benefits of a medical technology in terms of "quality-adjusted-life-years gained", combining life-years gained with an index representing the value of quality of life.[11]

This chapter summarises the way in which the benefits of liver transplantation can be estimated in the absence of firm comparative data. The same methods can of course be used to assess the benefits of other organ transplants.

Survival

To assess the benefits on survival, data concerning survival *with* transplantation have to be balanced against those for survival *without* transplantation. Four problems arise: 1) there are no data available from randomised studies, 2) the available data do not represent the current situation, 3) the differences between survival with and without transplantation may depend on patient characteristics and 4) the indications for transplantation may have changed over time.

SURVIVAL WITH TRANSPLANTATION

The probability of survival following transplantation can best be estimated on the basis of observed survival. Particularly if the probability of survival is likely to have changed over time, special care must be taken to ensure that such estimates represent the most recent situation. For liver transplant recipients, the overall probability of 5-year survival has improved steadily over the years, increasing from about 0.20 for patients treated between 1963 and 1979 to 0.64 for recipients treated between mid-1980 and 1988.[12,13] More recently an overall 3-year post-transplant survival probability of 0.682 was published, based on data from 8501 recipients undergoing transplantation between October 1987 and December 1991.[14] This suggests that the spectacular improvement noted up to 1988 did not continue. However, it is difficult to draw conclusions here. Survival after transplantation is affected not only by improvements in technique, follow up treatment or care, but also by a number of confounding factors such as age, race, primary liver disease and clinical status.[14-16] Hence, changes in the distribution of these factors, especially in the primary liver disease, may also change the overall

probability of survival. It is noted that in liver transplantation such confounding factors are probably more important than in heart or renal transplantation. Centre-to-centre variability in survival rates following liver transplantation can be reduced by 41% by the inclusion of indicators for patient-mix in the analysis.[17] In renal transplantation the inclusion of such indicators are least effective, reducing variability by only 12%.

To identify and correct for the simultaneous effects of changes in medical care and changes in the distribution of other confounding factors, both a large patient population and detailed data per patient are needed. When either of these is unavailable, correction is mostly limited to only one confounding factor at a time, e.g. diagnosis. Especially when evaluating the long term benefits of a technology on survival, such univariate analyses must lead to sub-optimal estimates. If numbers are too small to identify and correct for the effects of other confounding factors no correction can be made for e.g. the year of transplantation as a confounding factor. Such uncorrected survival curves do not necessarily give an accurate short-term survival probability for patients who had a transplantation in the more recent years of the period the curve was drawn for.

SURVIVAL WITHOUT TRANSPLANTATION

In the absence of data from a randomised clinical trial, a number of possibilities exists to estimate survival without transplantation (see Table 1). All methods have their own advantages and disadvantages and whichever method is chosen account has to be taken of the distribution of patient characteristics that affect the probability of survival, as with survival after transplantation.

Table 1 *Methods to estimate survival without transplantation*[18]

1. Quasi-experimental control group
2. Intervention delay group
3. Formal expert judgement
4. Historical control group
5. Prognostic modelling

A *quasi-experimental control group* could consist of patients who fulfilled entrance criteria but refused therapy. Before including a patient in such a control group, one has to be certain that the reason for refusing therapy is not related to the prognosis of the patient. Often such control groups are small, leaving estimates with large uncertainty margins.

Patients who are on the waiting list for the target technology could be included in an *intervention delay control group*. Such control groups are useful only if patients remain on the waiting list for some time and are treated in the sequence of entry on the list. If patients with the worst prognosis are given priority above patients with a better prognosis, a bias is introduced.

Estimates of survival probability based on *expert opinions* will be prone to subjectivity and uncertainty. Subjectivity, however, can also be seen as an advantage, since clinical judgement, which is often difficult to translate into objective factors, will probably be taken into account. The uncertainty margins surrounding this method can be narrowed by using a (relatively time consuming) Delphi method.[19] In this method a large number of experts are asked to give their opinion, anonymously and in writing. The results are analysed and reported back to the experts. They are then asked to restate their opinion, and the process continues until no further changes in opinion take place. However, the reliability and validity of the results remain difficult to test.

It is not always easy to find reliable *historical control groups*. Data gathered in the past will not have been recorded according to a systematic protocol, limiting the possibility of identifying whether patients would have fulfilled the current entry criteria. Additionally, if large changes in medical care have occurred, the results may not represent the current situation.

Sometimes a *prognostic model* can be used. With such a model a personal expected survival curve can be calculated for each patient at the time of transplantation, as if no transplantation had been performed. Individual survival curves can be combined into an aggregated curve, to produce an estimate that represents the average survival probability for the patient population that is to be assessed.[20]

For the estimation and identification of a reliable prognostic model, data on historical controls are most appropriate. Often such models are based on data from

patients with the whole range of disease stages. However, distinction should be made between estimation of the effect of confounders and estimation of survival. For estimation of the precise effect of the various confounders on survival, data may include patients who would not have been actual transplant candidates. For the precise estimation of survival probability, data from patients with the same stage of disease as transplant recipients should be used. Prognostic models combining both sources of data may introduce a false sense of reliability.

There are several examples of prognostic models for liver disease. Dickson and colleagues developed a Cox proportional hazards model to predict survival for patients with primary biliary cirrhosis,[21] often called the Mayo model. This model includes oedema, serum bilirubin and albumin levels, prothrombin time and age as variables. It has been validated several times, with good results for patient groups but inconsistent results for individuals.[22-24]

Other models for the estimation of survival in patients with primary biliary cirrhosis have been proposed by Christensen et al, Albers et al and Rydning et al.[25-27] Oellerich and colleagues and Christensen and co-workers produced models for patients with cirrhosis.[28-30] Models for patients with hepatitis B, primary sclerosing cholangitis and acute liver failure have also been published.[31-34]

SURVIVAL WITH AND WITHOUT TRANSPLANTATION: THE NUMBER OF LIFE-YEARS GAINED

The comparison between survival with and without transplantation is one of average results for the whole program. A change in the distribution of patient characteristics, or other time dependent changes in the liver transplantation programme, may influence both curves differently, leading to a different number of life-years gained for more recent transplant recipients. One of the underlying reasons for such a change in patient characteristics may be the availability of appropriate donor organs.[35] When more or fewer donor organs become available, the selection of patients may change. This may change the distribution of patient characteristics and therefore affect both survival curves (with and without transplantation) separately.

If the gain in survival is analysed for all patients with primary biliary cirrhosis who received transplants at Groningen University Hospital, the Netherlands,

between 1979 and 1990 (see chapter 4B), it is estimated that 3.5 years were gained by transplantation. If the more recent situation is described for these patients about 6.3 life years will be gained by transplantation instead of 3.5. This increase in the number of life-years gained is caused by an increase in the probability of survival with transplantation and by a decrease in the probability of survival without transplantation in the more recent transplant recipients. The decrease in survival probability without transplantation may have been caused by a shift in the timing of the transplantation towards a more advanced stage of disease.

Quality of life

Innovations may have an effect not only on survival but also on the quality of life. Especially for patients in the end-stage of a chronic disease, for whom a relatively long survival with a poor quality of life may be expected, the technology might affect quality of life more than survival. Here, as with survival, the effect of transplantation on quality of life is assessed by comparing quality of life *with* transplantation to that *without* transplantation. Moreover, as in the assessment of survival, the lack of randomised controlled double-blind trials limits empirical comparison.

The assessment of quality of life, leading to a concept that enables the calculation of quality adjusted life-years gained, consists of at least three steps: measurement, translation and valuation.[11] This section concentrates on the results from the measurements.

The individual patient, his or her family and the clinical practitioner will probably assess the benefits of the transplantation both in terms of survival and also by the subjective and objective health and well-being of the patient, the functioning of the transplant, the length of hospitalisation of the patient, and the occurrence of rejection and complications. In a formal assessment of the effects of transplantation on patient groups these latter effects are all comprised in the concept quality of life, to enable comparison with other health-care facilities. This quality of life is preferably measured by standard validated questionnaires that are sensitive to the changes expected in transplant recipients. Being sensitive means that detailed disease-specific questions have to be answered. To make comparisons possible, questions need to be answered that are not of particular importance to the disease being investigated. Questionnaires that combine both

aspects are rarely available and therefore two separate questionnaires are often used, one directed at the general health status, and the other at more disease-specific, or in this case transplant-specific, problems. In a number of circumstances it may be worthwhile to consider the use of a third questionnaire. For example when a value has to be assigned (normalised between zero and one) to observed health states, representing their relative weight between the worst and the best imaginable health state. In that case measurement instruments such as the EuroQol questionnaire or the Quality of Well-Being Scale,[36,37] which allow for an easy translation from measurement to valuation, will be helpful.

QUALITY OF LIFE WITH TRANSPLANTATION

Immediately after transplantation, patients are hospitalised for some time and most complications occur in this period. Consequently, the effects of transplantation on quality of life might be quite different in the short term than in the long term. To identify these differences, quality of life after transplantation must be measured at regular intervals and special care should be taken to avoid selective non-response. It may be that patients who are more severely ill do not respond because of their physical condition and this may affect the validity of results.[38] Selective non-response because of death should be taken into account in the interpretation but not in the calculation of the results.

In general, adult liver transplant recipients report good quality of life.[39-40] However, some authors report disturbances in neuropsychologic testing.[40] In a cross-sectional study in the UK it was reported that sleep was the most common problem and that problems with physical mobility, energy and social isolation decreased a year after transplantation.[4] The change in health status over time was also addressed in a longitudinal study in orthotopic liver transplant recipients in the Netherlands. Before transplantation patients showed psychological distress, a low Karnofsky index, many physical disturbances and a low level of experienced well-being.[38] Three months after transplantation a considerable improvement was observed, which was maintained for one year after transplantation.

QUALITY OF LIFE WITHOUT TRANSPLANTATION

Quality of life without transplantation is probably related to the stage of disease, and therefore data measuring quality of life without transplantation should be controlled for this. Unlike data concerning survival, no data can be made available about quality of life of historical control groups. This type of research is relatively new and data on quality of life have not been gathered consistently. Again, a second-best solution has to be found. The most practical one is to use quality of life experienced by patients while on the waiting list for transplantation as a proxy for their expected quality of life without transplantation. However, using this method, quality of life without transplantation will probably be overestimated, because it is assumed to remain steady; in reality there will probably be a time-dependent deterioration.

Benefits of transplantation

If the expected quality of life is taken into account, the number of Quality Adjusted Life-Years (QALYs) that might be gained by transplantation can be calculated. The conversion from life-years to QALYs is made in three steps.
First, the results of the measurements are translated into archetypal health states in five dimensions which represent mobility, self-care, activities of daily life, pain and depression-anxiety according to the EuroQol concept.[36]

Then a valuation on a scale of 0-1 is estimated for each health state. These estimations are based on valuations given by the general population to a large number of health states described in the same dimensions.[42]

Finally, the duration in each health state has to be multiplied with the value ascribed to each health state. Thereby the valuation is used as a correction factor on the two survival curves. For the patients with primary biliary cirrhosis who received transplants at Groningen University Hospital between 1979 and 1990 life years without transplantation have been multiplied with a factor 0.73 (chapter 4B). Additionally, the first three months after transplantation have also been multiplied by a factor 0.73, the next nine months by 0.84, the second year by 0.93 and the years thereafter by 0.87. The average number of QALYs gained by these patients is 3.45, compared with 3.52 life-years. Taking account of the enlarged differences for more recent transplant recipients the estimate is converted to 5.6 QALYs.

Discussion

Formal assessment of benefits of existing programmes is mainly carried out for economic evaluations. In such evaluations the other side of the coin of efficacy shows costs, and these are usually high in transplantation programmes, because the pretransplantation screening programme and life-long treatment with immunosuppressive drugs are taken into account.

Detailed information about the cost-effectiveness of medical technologies may support decisions on the allocation of scarce resources at several levels in the health-care sector.[43] On a national level, the decision regarding what proportion of a country's national income should be directed towards health care is at stake. Within health care, decisions on the reimbursement of new medical technologies or decisions regarding priorities towards certain diseases have to be made, as well as decisions on whether available resources for a certain disease should be spent on the treatment of end-stage patients or on prevention. For example, in the treatment of Hepatitis B money can be allocated to vaccination programmes or to transplant programmes, providing probably similar benefits in the long term. Likewise, it should be determined whether a new medical technology, such as liver transplantation, should be available for all indications. The outcome of these decisions may depend on whether short- or long-term effectiveness is sought. However, for all these decisions information is necessary. Only part of this information is formed by costs and benefits. Ethical considerations, legal restrictions and effects on equity should also be taken into account. Therefore, economic evaluation should be seen as a technique to offer information in a structured way, *alongside* other forms of information.

Within economic evaluation the precise assessment of benefits plays an important role, and we have outlined such an assessment with respect to liver transplantation. It should be stressed that the perspective chosen here has been one in which society's values are considered rather than individual expectations. Nevertheless, the one cannot be considered without the other and, as such, the same results may also help clinical decision-makers. Using similar techniques, groups of patients may be identified that benefit more (or less) by a service than other groups,[44] and such information may be considered - again with other information - in guidelines for the allocation of donor-organs in times of scarcity.[45] Additionally, these techniques may help to determine the optimal timing of

transplantation in the course of disease.[46] Such information has led to general recommendations for earlier referral of patients to transplant centres.[15,35,47]
Finally, of course, individual patients may be affected, directly or indirectly, by some of the above-mentioned decisions. However, because average expectations for patient groups are estimated, the results of these types of analyses will be of limited value to the individual patient.

In conclusion it can be stated that the assessment of the benefits of medical technology is crucial to providing the structured information that is needed to remove some of the uncertainties that always surround decision-making in health care. This chapter has given some methods and recommendations for such an assessment.

References

1 CPMP Working Party on Efficacy of Medicinal Products. EEC Note for guidance: Good clinical practice for trials on medicinal products in the European Community. Pharmacology and Toxicology 1990; 67: 361-372

2 Starzl TE, Todo S, Tzakis AG, et al. Liver transplantation: an unfinished product. Transplantation Proceedings 1989; 21: 2197-2000

3 Starzl TE, Iwatsuki S, Thiel DH van, et al. Evolution of liver transplantation. Hepatology 1982; 2: 614-636.

4 Austen WG, Cosimi AB. Editorial retrospective: Heart transplantation after 16 years. New England Journal of Medicine 1984; 311: 1436-1438

5 McCaughan GW, McDonald J, Bell R, et al. Improved survival after liver transplantation for primary biliary cirrhosis. Transplantation Proceedings 1989; 21: 3788-3789

6 Anonymous. Non-randomised controls and assessment of survival in liver transplantation [editorial]. Lancet 1990; 335: 509-510

7 Sheil AGR. Quandaries and controversies in liver transplantation. Transplantation Proceedings 1992; 24: 2375-2378

8 Drummond M, Rutten F, Brenna A, et al. Economic evaluation of pharmaceuticals. A European perspective. PharmacoEconomics 1993; 4: 173-186

9 Williams A. Economics of coronary artery bypass grafting. British Medical Journal 1985; 291: 326-329

10 Elixhauser A, Luce BR, Taylor WR, Reblando J. Health care CBA/CEA: an update on the growth and composition of the literature. Medical Care 1993; 31: JS 1-11, JS 18-149

11 Torrance GW. Measurement of Health State Utilities for economic appraisal. Journal of Health Economics 1986; 5: 1-30
12 Iwatsuki S, Starzl TE, Todo S, et al. Experience in 1,000 liver transplants under cyclosporine-steroid therapy: a survival report. Transplantation Proceedings 1988; 20: 498-504
13 Starzl TE, Demetris AJ, Thiel D van. Liver transplantation (second of two parts). The New England Journal of Medicine 1989; 321: 1092-1099
14 Belle H, Detre KM. Report from the Pitt-UNOS liver transplant Registry. Transplantation Proceedings 1993; 25: 1137-1142
15 Gunson BK, Buckels JA, Elias E, et al. Liver transplantation for primary biliary cirrhosis: the effect of disease stage at referral upon outcome after liver replacement. Transplantation Proceedings 1990; 22: 1503-1504
16 Mora NP, Klintmalm GB, Solomon H, et al. Survival after liver transplantation in 300 consecutive patients: the influence of age, clinical status, and pretransplant disease. Transplantation Proceedings 1992; 24: 156-157
17 Hunsicker LG, Edwards EB, Breen TJ, Daily OP. Effect of center size and patient-mix covariates on transplant center-specific patient and graft survival in the United States. Transplantation Proceedings 1993; 25: 1318-1320
18 Bonsel GJ. Methods of medical technology assessment with an application to liver transplantation. PhD thesis, Erasmus University Rotterdam, 1991.
19 Fink A, Kosecoff J, Chassin M,, Brook RH. Consensus methods: characteristics and guidelines for use. American Journal of Public Health 1984; 74: 979-983
20 Thomsen BL, Keiding N, Altman DG. A note on the calculation of expected survival, illustrated by the survival of liver transplant patients. Statistics in Medicine 1991; 10: 733-738
21 Dickson ER, Grambsch PM, Fleming TR, et al. Prognosis in primary biliary cirrhosis: model for decision making. Hepatology 1989; 10: 1-7
22 Grambsch PM, Dickson ER, Kaplan M, et al. Extramural cross-validation of the Mayo primary biliary cirrhosis survival model establishes its generalizability. Hepatology 1989; 10: 846-850
23 Bonsel GJ, Klompmaker IJ, van't Veer F, et al. Use of prognostic models for assessment of value of liver transplantation in primary biliary cirrhosis. Lancet 1990; 335: 493-497
24 Jeffrey GP, Hoffman NE, Reed WD. Validation of prognostic models in primary biliary cirrhosis. Australian and New Zealand Journal of Medicine 1990; 20: 107-110
25 Christensen E, Neuberger J, Crowe J, et al. Beneficial effect of azathioprine and prediction of prognosis in primary biliary cirrhosis. Final results of an international trial. Gastroenterology 1985; 89: 1084-1091.
26 Albers I, Hartmann H, Bircher J, Creutzfeldt W. Superiority of the Child-Pugh classification to quantitative liver function tests for assessing prognosis of liver transplantation. Scandinavian Journal of Gastroenterology 1989; 24: 269-276

27 Rydning A, Schrumpf E, Abdelnoor M, et al. Factors of prognostic importance in Primary Biliary Cirrhosis. Scandinavian Journal of Gastroenterology 1990; 25: 119-126
28 Oellerich M, Burdelski M, Lautz HU, et al. Assessment of pretransplant prognosis in patients with cirrhosis. Transplantation 1991; 51: 801-806
29 Oellerich M, Hartmann H, Ringe B, et al. Assessment of prognosis in transplant candidates by use of the Pugh-MEGX score. Transplantation Proceedings 1993; 25: 1116-1119
30 Christensen E, Schlichting P, Andersen PK, et al. Updating prognosis and therapeutic effect evaluation in cirrhosis with Cox' multiple regression model for time dependant variables. Scandinavian Journal of Gastroenterology 1986; 21: 163-174
31 Bernuau J, Goudeau A, Poynard T, et al. Multivariate analysis of prognostic factors in fulminant hepatitis B. Hepatology 1986; 6: 648-651.
32 Wiesner RH, Grambsch PM, Dickson ER, et al. Primary sclerosing cholangitis: natural history, prognostic factors and survival analysis. Hepatology 1989; 10: 430-436
33 Farrant JM, Hayllar KM, Wilkinson ML, et al. Natural history and prognostic variables in primary sclerosing cholangitis. Gastroenterology 1991; 100: 1710-1717
34 Christensen E, Bremmelgaard A, Bahnsen M, et al. Prediction of fatality in fulminant hepatic failure. Scandinavian Journal of Gastroenterology 1984; 19: 90-96
35 Lynch SV, Balderson G, Armstrong G, et al. Mortality in patients on liver transplant waiting lists: an avoidable tragedy. Transplantation Proceedings 1989; 21: 3794-3796
36 The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy 1990; 16: 199-208
37 Patrick DL, Bush JW, Chen MM. Towards an operational definition of health. Journal of health and social behaviour 1973; 14: 6-23
38 Bonsel GJ, Essink-Bot ML, Klompmaker IJ, Slooff MJH. Assessment of the quality of life before and following liver transplantation. Transplantation 1992; 53: 796-800
39 Foley TC, Davis CP, Conway PA. Liver transplant recipients-self-report of symptom frequency, symptom distress, quality of life. Transplantation Proceedings 1989; 21: 2417-2418
40 Wolcott D, Norquist G, Busuttil R. Cognitive function and quality of life in adult liver transplant recipients. Transplantation Proceedings 1989; 21: 3563
41 Lowe D, O'Grady JG, McEwen J, Williams R. Quality of life following liver transplantation: a preliminary report. Journal of the Royal College of Physicians of London 1990; 24: 43-46
42 Essink-Bot ML, Bonsel GJ, Maas PJ van der. Valuation of health states by the general public: feasibility of a standardized measurement procedure. Social Science and Medicine 1990; 31: 1201-1206
43 Krueger H. Economic analysis of solid organ transplantation: a review for policy makers. Health Policy 1989; 13: 1-17.

44 Foster WR, Burton BT. Technology assessment applied to liver transplantation in adults. International Journal for Technology Assessment in Health Care 1989; 5: 173-182.

45 Donovan JP, Zetterman RK, Burnett DA, Sorrell MF. Preoperative evaluation, preparation, and timing of orthotopic liver transplantation in the adult. Seminars in Liver Disease 1989; 9: 168-175

46 Wiesner RH, Porayko MK, Dickson ER, et al. Selection and timing of liver transplantation in Primary Biliary Cirrhosis and Primary Sclerosing Cholangitis. Hepatology 1992; 16: 1290-1299

47 Neuberger JM, Gunson BK, Buckels JA, et al. Referral of patients with primary biliary cirrhosis for liver transplantation. Gut 1990; 31: 1069-1072

A SURVEY OF THE COSTS AND EFFECTS OF LIVER TRANSPLANTATION IN THE LONG TERM; THE LIVER TRANSPLANTATION PROGRAMME OF GRONINGEN UNIVERSITY HOSPITAL: 1979-1991*

Summary

This chapter presents an example of the use of modelling in a cost-effectiveness analysis where no contemporary control group was available. A medical technology assessment of the liver transplantation programme of the Groningen University Hospital, which was commissioned by the Dutch National Health Insurance Funds Council, is discussed. The results of all 152 liver transplantations performed between 1979 and November 1990 were analysed. The main objective of the study was to evaluate the long term effects of liver transplantation. Five years post-transplantation 59% of the adult patients were still alive. The survival probability highly depended on the primary liver disease. Cox regression was used to model survival without liver transplantation. Mortality and morbidity mainly occurred within the first year post-transplantation. Hereafter the prospects for the patients were excellent, both with regard to the chances of survival and to the quality of life. The costs of liver transplantation could be calculated only for patients with primary biliary cirrhosis and other forms of biliary cirrhosis, not caused by hepatitis B infection or alcohol abuse. For these patients the costs of a liver transplantation were estimated at DFL 263,000 (with a 10% margin), 10 years of follow-up included and corrected for the costs that would have been made for the treatment of the liver disease. For this population of patients a liver transplantation costs between DFL 64,000 and DFL 79,000 per life year gained, 10 years of follow-up included. The major part of the costs is taken up by the necessary prescription of cyclosporin. The need for liver transplantation in the Netherlands was estimated at between 35 and 126 transplantations per year,

* Based on: Michel BC, Bonsel GJ, Stouthard MEA, McDonnell J, Habbema JDF. Kosten-effectiviteitsanalyse van levertransplantatie op lange termijn; het Groningse levertransplantatieprogramma 1979-1991. Nederlands Tijdschrift voor Geneeskunde 1993; 137(19): 963-969

depending on indications, contra-indications, referral and the percentage of re-transplantations.

Introduction

The first liver transplantation in man was performed by Starzl in the US in 1963. The patient survived for 22 days.[1] Since 1979 liver transplantations have also been performed in the Netherlands. In the beginning only at Groningen University Hospital, but since 1986 also at Rotterdam University Hospital. Initially, the liver transplantation programme in Groningen was financed by the Dutch Health Insurance Funds Council on a temporary basis. In 1985 this council commissioned an independent assessment of the liver transplantation programme of Groningen University Hospital to Erasmus University in Rotterdam. They aimed to use the results of this assessment for the determination of the place of liver transplantation in Dutch health insurance. Such an assessment of, primarily, the costs and effects of a medical technology, aimed at informed political decision making, is often called a "Medical Technology Assessment" (MTA). In 1988 this assessment was completed.[2,3] Because of the small number of patients with a long follow-up no conclusion on the long-term effects of liver transplantation could be drawn. Partly because of that, the former Secretary of State for Public Health and the Health Insurance Funds Council together decided to postpone the decision and to instigate a second study focusing on the long-term effects of liver transplantation. The main results of this second study, completed in January 1992, are presented in this chapter.[4]

Patients and methods

For every patient referred to Groningen University Hospital for liver transplantation the reason of referral, the time spent in the separate phases of the transplant programme, the date patients were excluded from the programme or the date of transplantation, and the date of death were registered. For every liver transplant in an adult patient the physicians at Groningen University Hospital collected demographic data on the recipient, and data concerning the pre-operative condition of the recipient, the surgery and the donor. Following transplantation, data on the recipient were collected at three months and at 1, 3, 5, and 10 years.

Also, the date and character of each episode of illness, and the ward and length of each stay in hospital or nursing-home following the transplantation were registered.

In categorising the diagnoses the group "other cirrhoses" was formed. In this category auto-immune cirrhosis, cryptogenic cirrhosis and other forms of cirrhosis were included, but not cirrhosis after viral hepatitis or alcoholic cirrhosis.
The number of episodes of illness was expressed per patient-year because not all patients were at risk during the whole period, due to death, re-transplantation or censuring. In presenting changes over time, period I (1979 to September 1987) was compared to period II, the period of the second study (October 1987 to November 1990).

For the assessment of quality of life adult patients were asked to complete questionnaires, presented by personal computer, in the period July 1987 to March 1991.[5] During their annual in-hospital check-up patients were asked to sit behind the personal computer in a sequestered place. The computer programme was constructed in such a way that patients could complete the questionnaire independently after a short instruction by the staff-members. The questionnaire was completed in writing only if patients were too ill or if there were problems in the organisation of a computer session. Mostly, existent validated questionnaires were used. Some questionnaires were developed for this study.

ANALYSES AND MODELS

For the analysis of the probability of survival the method developed by Kaplan and Meier was used.[6,7] Determinants of prognosis were evaluated by Cox proportional hazards model.[8] To calculate the number of life-years that might be gained by transplantation the expected survival without transplantation was estimated by modelling techniques.[9,10] Survival without transplantation for patients with "other cirrhosis" was predicted with a model that was based on data from patients at the Hospital Clinic i Provincial de Barcelona who didn't receive a transplant (yet).[11] To predict survival without transplantation for patients with primary biliary cirrhosis a model developed in the Mayo Clinic in the US was used.[12]

The number of life-years that might be gained by liver transplantation was corrected for quality of life and expressed in "quality adjusted life-years" (QALYs).[13]

COSTS

For the cost-calculation data from the first technology assessment were used.[14] Adaptations were made for changes that could be deduced from the current data collection. New data concerning the probability of survival, the number of hospitalisations and the length of stay in hospital, the proportion of re-transplantations and the number of patients in each pre-transplantation phase were available. Data collection for the second MTA study was completed at November 1, 1990.

Costs and effects were discounted by 5% annually: this means that life-years that are gained in the short term are valued higher than those gained in the long term.[15] In the calculations of the costs, margins of uncertainty of 10% were used.

Results

Between 1979 and November 1990 in total 745 patients were referred to Groningen University Hospital for transplantation; 19% of these patients received a transplant in this period. The largest number of referrals was seen in 1983. Over the years the number of drop-outs in the separate phases decreased (table 1). In 1983 9.9% of the referred patients received a transplant, in 1990 this had increased to 43%. In period I still 47 of the referred patients with primary biliary cirrhosis or "other cirrhosis" (41%) were thought to be unfit for transplantation because of inferior liver function, in period II this happened only once (6%).

Between 1979 and November 1990 152 liver transplants were received by 134 patients, 33 of those were children. Eighteen re-transplantations were performed. The main indication for transplantation was primary biliary cirrhosis (table 2). Five years following transplantation 59% of the adult recipients were still alive. The main mortality was seen in the peri-operative period and within the first year. Patients who survived the first year had a probability of 88% to survive 5 years.

Table 1 *Annual inflow of patients (adults and children) in separate phases of the liver transplantation programme at Groningen University Hospital, 1977 to end 1990.*

	R-phase[a]	PC-phase[b]	A-phase[c]	GL-phase[d]	LTX-phase[e]
1977	2	1	1	1	0
1978	9	1	0	0	1
1979	21	7	4	4	4
1980	23	7	5	5	3
1981	43	24	5	5	5
1982	84	33	19	11	11
1983	101	49	25	9	10
1984	91	32	27	8	6
1985	78	24	16	11	7
1986	59	27	27	13	11
1987	57	27	18	18	15
1988	57	35	31	20	14
1989	60	39	31	22	27
1990	60	28	24	34	26
Total	**745**				**140**[‡]

[a] *R-phase = referral: patients are referred to Groningen University Hospital for liver transplantation*

[b] *PC-phase = possible candidate: transplantation is being considered, in-hospital evaluation.*

[c] *A-phase = acceptance: patients are accepted for transplantation and await the right timing of transplantation in their course of illness.*

[d] *GL-phase = green light: patients are on the Eurotransplant waiting-list waiting for an appropriate donor.*

[e] *LTX-phase = liver transplantation: patients receive a (primary) liver transplant*

[‡] 134 primary transplantations + 5 transplantations that were performed in 1990 after November 1st + 1 transplantation performed in London in 1978

The primary liver disease had a strong influence on survival following liver transplantation. Patients with biliary atresia (mainly children) had the best prognosis, patients with "other cirrhosis" the worst, mainly because of a large peri-operative mortality (figure 1a). The age of the patient and the severity of the liver disease, expressed in the Child-Pugh score,[16] only had a small influence on post-transplant survival.

Table 2 *Demographic and medical data on adult patients at the time of their first transplantation. Comparison of period I (1979 to October 1987) and period II (October 1987 to November 1990).*

	Period I	Period II
Number of patients	59	42
Diagnosis: Malignant tumour (%)	2 (3%)	1 (2%)
Primary biliary cirrhosis	24 (41%)	11 (26%)
Other cirrhosis	22 (37%)	10 (24%)
Primary sclerosing cholangitis	3 (5%)	5 (12%)
Other diagnoses	8 (14%)	15 (36%)
Gender: number of females (%)	43 (73%)	27 (64%)
Age: average in years (SD)	40.8 (12.3)	42.6 (11.9)
Duration of illness: average in years (SD)	7.2 (4.0)	7.4 (5.8)
Hepatorenal syndrome: number (%)	26 (44%)	33 (79%)
Portal hypertension: number (%)	51 (86%)	35 (83%)
Child-Pugh class[16]: class A (%)	16 (27%)	5 (12%)
class B	32 (54%)	16 (38%)
class C	11 (19%)	21 (50%)
Karnofsky-score: average (SD)	63.1 (17.7)	55.5 (25.8)
ALAT/SGOT (0 - 40 U/L)	124.8 (81.6)	153.7 (143.5)
Bilirubin total (3 - 26 mmol/L)	128.5 (160.5)	195.0 (196.6)
Alkaline phosphatase (13 - 120 U/L)	600.7 (650.5)	473.2 (428.3)
γ-Glutamyl transpeptidase (0 - 65 U/L)	322.5 (366.3)	230.0 (296.5)
Albumin (34 - 47 g/L)	33.3 (6.4)	29.1 (6.6)

Figure 1a *Survival after liver transplantation in 134 patients (adults and children) in the period 1979 - November 1990 by diagnosis (A: biliary atresia, n = 16 all children; B = other diagnoses n = 45; C = primary biliary cirrhosis n = 35; D = other cirrhosis n = 38, including 6 children)*

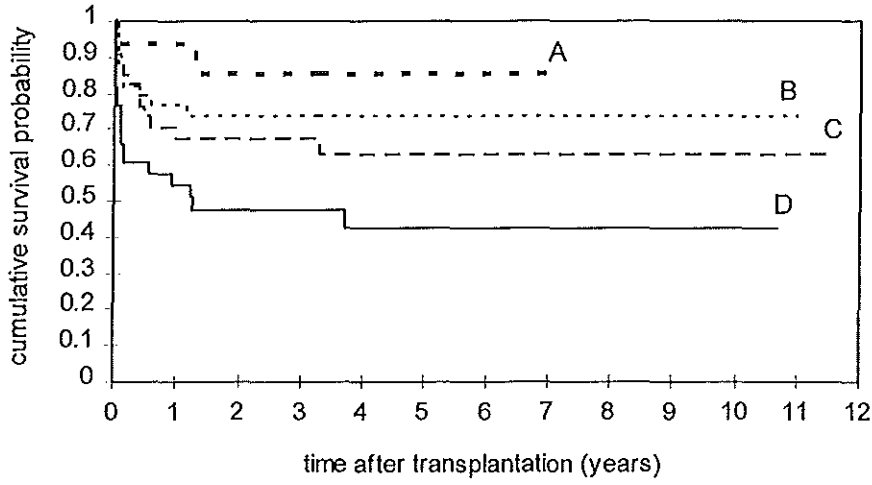
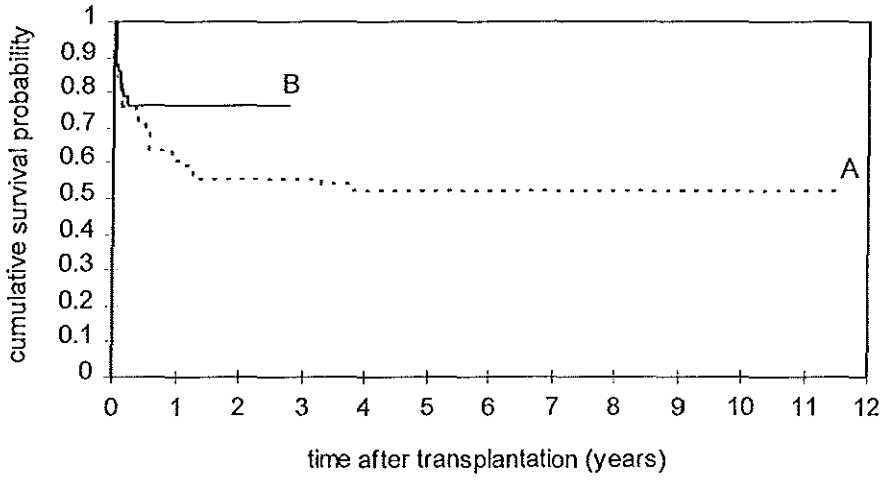


Figure 1b *Survival after liver transplantation in 101 adult patients, comparing period I (A) to period II (B)*



During the liver transplantation programme a clear improvement of the probability of survival was seen for adult patients (figure 1b). A strongly improved probability of survival from 3 months post-transplantation onwards was seen in period II. Peri-operative mortality and mortality within the first 3 months did not change.

The proportion of re-transplantations to primary transplantations was identical in both periods. But the time between the first and the second transplantation clearly decreased, from, on average, 11.6 months in the first period to 1.6 months in the second. In many patients the transplant functioned well from three months post-transplantation onwards. Liver functions improved until 5 years post-transplantation and reached normal values in a large fraction of the patients (table 3).

Table 3 *Function of the liver transplant, liver cell decay and cholestasis in adult patients in the long term. Median value (percentage of patients outside normal range).*

Liver functions (normal range)	Pre-operative n = 109[1]	3 months after surgery n = 78[2]	5 years after surgery n = 17[3]
SGOT/ALAT (0-40 U/L)	116.0 (86%)	35.0 (43%)	24.0 (29%)
γ-GT (0-65 U/L)	183.0 (76%)	230.0 (74%)	21.0 (29%)
LDH (114-235 U/L)	252.0 (61%)	279.0 (73%)	231.0 (41%)
Alkaline phosphatase (13-120 U/L)	344.0 (87%)	152.5 (62%)	60.0 (29%)
Bilirubin total (3-26 mmol/L)	68.0 (73%)	21.5 (41%)	16.0 (12%)
Albumin (34-47 g/L)	32.0 (62%)	38.0 (21%)	41.0 (0%)
Protrombin time (0-16 sec)	15.0 (31%)	12.8 (4%)	13.0 (0%)
Antithrombin III (>80%)	72.0 (60%)	117.0 (4%)	103.0 (12%)

[1] *except protrombin time n = 110; albumin n = 111*
[2] *except protrombin time n = 54; antithrombin III n = 57; LDH n = 77*
[3] *except protrombin time n = 16*

On average patients experienced 10 episodes of illness per patient-year during the first year post-transplantation. After the first year this number decreased to, on average, 1.2 episodes of illness per patient-year in the second year. In later

years a further decrease in the average number of episodes was seen to 0.3 episodes in the ninth year (in 10 patients).

Patients spent much time in hospital during the first year following transplantation (on average 2.4 months per patient), afterwards this decreased to on average 3.2 days per patient in the sixth year. In period II patients, on average, spent more time on artificial respiration and in the Intensive Care unit than during period I.

QUALITY OF LIFE AND LIFE-YEARS GAINED

Before transplantation patients reported a moderate quality of life (table 4) The main complaints were related to the liver disease, frequent sleep disorders were reported and there were restrictions in mobility and energy. Candidates for transplantation were (on average) somewhat frightened and rather depressed. But they still had a fair ability to take care of themselves. From 3 months after transplantation the physical, psychological and subjective health-state improved, and one year after transplantation a stable situation had established that was almost identical to the level of an average population of healthy adults, except for the number of hours that was daily spent in school or at a paid or unpaid job.

The average number of life-years gained per patient by liver transplantation was estimated by comparing the observed survival *with* transplantation to the estimated survival *without* for patients with primary biliary cirrhosis and "other cirrhosis" separately, figure 2. With this method it was calculated that a patient with primary biliary cirrhosis gained on average 3.2 years in the first 10 years after transplantation (with discounting). A patient with "other cirrhosis" gained 0.9 years during the same follow-up. This difference is mainly caused by a much lower survival following transplantation for patients with "other cirrhosis". After correction for quality of life the average number of life-years (QALYs) that was gained by patients with primary biliary cirrhosis and "other cirrhosis" together was 3.5 years with a time-limit of 10 years and 6.9 years within a follow-up of 25 years.

Table 4 Quality of life in 70 patients before and following liver transplantation in the period 1987 -1991.[a]

		Questionnaire completed[b]			
	Reference values[c]	pre-operative (SD)	after 3 months	after 1 year	after 5 years
Questionnaire[d]		(n = 43)	(n = 31)	(n = 25)	(n = 7)
Daily activities (ADL)[16]	1-10 (≥9)	8.3 (1.9)	8.8	9.4	9.9
Hours work/school daily	0-12 (±8)	1.9 (2.9)	1.4	3.4	4.4
Karnofsky[17]	0-100 (≥90)	61.4 (19.6)	69.4	89.2	87.7
Anxiety (STAI)[18]	80-20 (≤37)	43.0 (10.8)	35.2	31.7	30.3
Depression (SDS)[19,20]	100-25 (≤35)	43.4 (8.8)	35.7	32.6	36.6
Well-being (IWB)[21]	2.1-14.7 (>12)	8.9 (3.2)	12.6	12.6	13.6
NHP-DA[22,23]					
-mobility	100-0 (≤15)	36.2 (28.5)	27.2	15.2	17.1
-pain	100-0 (≤15)	18.9 (26.8)	11.8	5.2	11.7
-energy	100-0 (≤15)	67.9 (39.7)	20.5	4.9	8.7
-sleep	100-0 (≤15)	42.5 (35.9)	16.8	13.5	5.9
-social isolation	100-0 (≤15)	18.8 (23.1)	7.5	8.8	5.5
-emotional reaction	100-0 (≤15)	21.2 (22.7)	8.0	4.4	2.9

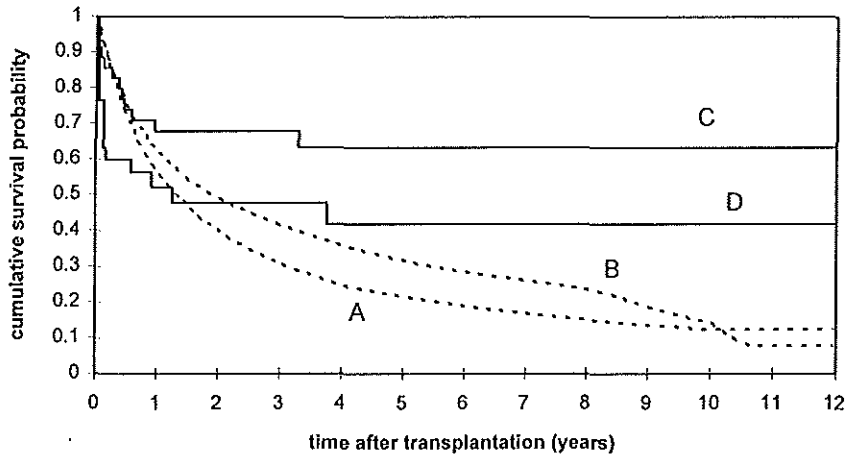
[a] *Data were collected in the period July 1987 - March 1991: 182 questionnaires were completed by 70 patients, 11 of these patients had not received a liver transplant by March 1991 yet; 23 other questionnaires were not completed, 2 because of refusal by the patient, 15 because of problems in organising a session for the patient, 6 because the patient was too ill.*

[b] *Standard deviations are only mentioned for pre-operative values. For later scores the standard deviation was identical or smaller.*

[c] *Range of possible scores (reference values normal population); worst score left, best score right*

[d] *ADL = activities of daily life; STAI = state trait anxiety inventory; SDS = self rating depression scale; IWB = index of well-being; NHP-DA = Nottingham health profile, Dutch adaptation*

Figure 2 *Probability of survival without transplantation predicted for patients with primary biliary cirrhosis (A) and "other cirrhosis" (B) compared with the observed survival following liver transplantation (respectively C and D)*



COSTS

Costs of liver transplantation, including 10 years of after-care, were estimated at DFL 319,000 for patients with primary biliary cirrhosis and "other cirrhosis" (with a margin of uncertainty of 10%). Opposite these costs were the savings due to the disappearance of the costs that would have to be made for treatment of the liver disease, approximately DFL 56,000. The differential costs were DFL 263,000 per patient (with a margin of uncertainty, discounted), including 10 years of after-care.

Per life-year gained liver transplantation costs between DFL 64,000 and DFL 79,000 including 10 years after-care for patients with primary biliary cirrhosis and "other cirrhosis". After correction for quality of life, one QALY costs between DFL 69,000 and DFL 84,000 with a time-limit of 10 years. Costs of cyclosporin still play an important role. If the price or the dosage of cyclosporin would be halved, cost-effectiveness would improve to between DFL 56,000 and DFL 69,000 per QALY gained for patients with primary biliary cirrhosis or "other cirrhosis".

THE NEED FOR LIVER TRANSPLANTATION IN THE FUTURE

The need for liver transplants in the future depends on the contra-indications, the probability of a re-transplantation, the indications and the referral-policy. It is not to be expected that contra-indications and the probability of a re-transplantation will change drastically in the coming years. Within the indications, hepatitis B and, to a lesser extend, acute liver failure might cause an increase in the need for transplant capacity. Based on the incidence of both diseases it is to be expected that minimally 4 and maximally 47 transplants a year will be performed for these indications.

If liver diseases are evenly spread over the Netherlands, referral to Groningen University Hospital has not been optimal so far. Certain districts were strongly represented, like Friesland (5 transplants / million inhabitants in 1988) and Drenthe (3.8 transplants / million inhabitants in 1988); other districts were badly represented like Zeeland (0.9 transplants / million inhabitants in 1988). A more evenly spread referral policy could cause an increase in the number of transplantations, if there is no over-treatment in the northerly districts.

Based on these assumptions it is estimated that annually between 35 and 126 liver transplants will be needed.

Discussion

Long-term results of liver transplantation appear to be very good. If a patient survives the first year following transplantation his or her perspectives are excellent.

Over the last years clear changes have taken place in the Groningen University Hospital liver transplantation programme. Firstly, less patients drop out of the programme in the pre-transplantation phases. This could indicate an improved referral to Groningen University Hospital or more liberal indications used by Groningen University Hospital, with less weight on contra-indications. Secondly, the probability of survival following hospitalisation is better than before. Because this improvement only shows after 3 months, an improved after-care and treatment with cyclosporin (since the end of 1985) probably play an important role in this.

Other centres also saw an improvement in survival after the introduction of cyclosporin.[25] Thirdly, a change in the indications for liver transplantation took place. Hepatitis B and acute liver failure were added to the indications on a limited scale.

It can be expected that the number of liver transplantations in the Netherlands will increase somewhat in the future. The number of liver transplants, with 2.7 transplants per million inhabitants in 1990, is lower than in the rest of Europe. In West-Germany, Austria and Belgium this number is 5.0, 10.5 and 14.0 respectively.[26,27] In the calculation of these numbers no account is taken of the sometimes high percentage of patients that are referred to specialised centres from abroad: such flows of patients will influence the number of transplants per million inhabitants of both the referring and the receiving country. Other explanations for these differences can be found in epidemiological differences (hepatitis B, alcohol), the policy regarding the acceptance of patients, especially with malignant tumours or hepatitis B, and in age-limits.

In 1990 73 donor livers became available for transplantation in the Netherlands, whereas in the future between 35 and 126 livers will be needed.[28] It is not unthinkable that a shortage in donor-livers will occur if an increase in need is combined with a stable supply.

The costs of liver transplantation could be calculated with a higher precision in the study presented here than in the first study, because more data concerning long-term effects were available and because Groningen University Hospital clearly passed the starting period. In the first study costs per life-year gained were estimated at between DFL 47,000 and DFL 133,000. Now this range could be adjusted to between DFL 64,000 and DFL 79,000 per life-year gained.

In recent years more voices have been heard demanding that considerations concerning costs and effects will be allowed to play an important role in the judgement of new medical technology. The committee "Choices in Care" recommends the use of, amongst others, efficacy and efficiency in the decision whether or not to include certain technology in baseline insurance.[29] The report "Medical actions at cross-roads" recommends the introduction of new technology on a limited scale, followed by an evaluation in clinical practice before a large scale introduction is done.[30] This procedure was followed in liver transplantation. It was introduced (and financed) on a limited scale and the decision on whether or not to

include it in baseline insurance was postponed until the inventory of its costs and effects was completed. Meanwhile the Health Insurance Funds Council has offered a recommendation, based on the MTA-report, to the Secretary of State for Public Health. A decision about the inclusion of liver transplantation in the national sickness fund is expected soon.

References

1 Starzl TE, Marchioro TL, Kaulla KN von, et al. Homotransplantation of the liver in humans. Surg Gynecol Obstet 1963; 117(6): 659-676.

2 Bonsel GJ, Habbema JDF, Bot ML, Veer F van 't, Charro FT de, Maas PJ van der. Een 'technology assessment' van levertransplantatie: een onderzoek van het Groningse levertransplantatieprogramma 1977 - 1987. Ned Tijdschr Geneeskd 1989; 133 (28): 1406-1414.

3 Bonsel GJ, Klompmaker IJ, Essink-Bot ML, Habbema JDF, Slooff MJH. Cost-effectiveness analysis of the Dutch liver transplantation programme. Transplantation Proceedings 1990; 22: 1481-1484.

4 Michel BC, Bonsel GJ, Stouthard MEA, et al. Levertransplantatie: de effectiviteit op lange termijn. Rotterdam: Instituut Maatschappelijke Gezondheidszorg. Erasmus Universiteit, 1992.

5 Bonsel GJ, Essink-Bot ML, Klompmaker IJ, Slooff MJH. Quality of life after liver transplantation: first results. Transplantation 1992; 53(4): 796-800.

6 Kaplan EL, Meier P. Non-parametric estimation from incomplete observations. J Am Stat Ass 1974; 69: 74-79.

7 Houwelingen JC van. De statistiek van overlevingscurven. Ned Tijdschr Geneeskund 1991; 135 (38): 1735-1744.

8 Christensen E. Multivariate survival analysis using Cox's regression model. Hepatology 1987; 7 (6): 1346-1358.

9 Bonsel GJ, Klompmaker IJ, van 't Veer F, Habbema JDF, Slooff MJH. Use of prognostic models for assessment of liver transplantation in primary biliary cirrhosis. Lancet 1990; 335: 493-497.

10 Editorial. Non-randomised controls and assessment of survival in liver transplantation. Lancet 1990; 335: 509-510.

11 Llach J, Ginès P, Arroyo V, et al. Prognostic value of arterial pressure, endogenous vasoactive systems, and renal function in cirrhotic patients admitted to the hospital for the treatment of ascites. Gastroenterology 1988; 94: 482-487.

12 Dickson ER, Grambsch PM, Fleming TR, et al. Prognosis in PBC: model for decision making. Hepatology 1989; 10: 1-7.

13 Torrance GW. Measurement of health utilities for economic appraisal. A review. J Health Econ 1986; 5: 1-30.
14 Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1978.
15 Pugh RNH, Murray-Lyon IM, Dawson JL, et al. Transsection of the oesophagus for bleeding oesophageal varices. Brit J Surg 1973; 60(8): 646-649.
16 Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged: The index of ADL: a standardized measure for Biological and Psychosocial functioning. JAMA 1963; 185: 94-99.
17 Grieco A, Long CJ. Investigation of the Karnofsky performance status as a measure of quality of life. Health Psychol 1984 ; 3(2) :129-142.
18 Ploeg HM van der. Validiteit van de ZelfBeoordelingsVragenlijst. Ned Tijdschr Psychol 1980; 35: 243-249.
19 Zung WWK. From art to science: The diagnosis and treatment of depression. Arch. Gen. Psychiatry 1973; 20: 328-337.
20 Zitman FG, Griez EJL, Hooijer Chr. Standaardisatie van depressie vragenlijsten. Ned Tijdschr Psychiatr 1989; 31: 114-135.
21 Campbell A, Converse PE, Rodgers WL. The quality of American life. New York: Russell Sage Foundation, 1976.
22 Hunt SM, McEwen J, McKenna SP. Measuring health status. London: Croom Helm, 1986.
23 Essink-Bot ML, Agt HME van, Bonsel GJ. NHP of SIP: een vergelijkend onderzoek onder chronisch zieken. Tijdschr. Soc. Gen. 1992; 70: 152-159.
24 Iwatsuki S, Starzl TE, Gordon RD, et al. Experience in 1,000 liver transplants under cyclosporine-steroid therapy: a survival report. Transplantation Proceedings 1988; 20 (1 Suppl 1): 498-504.
25 Eurotransplant Foundation. Annual report 1990. Ed. by Guido Persijn, Bernard Cohen. Leiden.
26 Centraal Bureau voor de Statistiek. Maandstatistiek van de Bevolking: september 1990. Voorburg: CBS/SDU, 1990.
27 Persijn G, Vergelijkend overzicht van postmortale orgaandonatie in Nederland. Nederlands Tijdschrift voor Geneeskunde 1991; 135 (36): 1652-53.
28 Ministerie van Welzijn, Volksgezondheid en Cultuur: Commissie Keuzen in de zorg. Kiezen en delen. Rijswijk: Ministerie van WVC, 1991.
29 Gezondheidsraad: Beraadsgroep Geneeskunde. Medisch handelen op een tweesprong. Den Haag: Gezondheidsraad, 1991; publicatie nr 1991/23.

ECONOMIC ASPECTS OF TREATMENT WITH CAPTOPRIL FOR PATIENTS WITH ASYMPTOMATIC LEFT VENTRICULAR DYSFUNCTION IN THE NETHERLANDS*

Summary

The objective of the study presented in this chapter is to estimate costs and effects of preventive treatment with captopril compared to the current treatment policy in patients with asymptomatic left ventricular dysfunction after a myocardial infarction. Estimates of effects are based on the results from the SAVE-trial. Costs are estimated on the basis of current treatment patterns in four Dutch hospitals. All knowledge is incorporated in a mathematical model extrapolating the SAVE-results to 20 years and complementing the SAVE-results with cost data. Captopril treatment is expected to increase survival at certain costs. The average additional costs per patient are estimated at DFL 2,491 in 4 years and at DFL 8,723 in 20 years of treatment. Costs per additional survivor after four years are estimated at DFL 69,126. After extrapolation of the results of the SAVE trial to 20 years, costs per life year gained can be estimated at DFL 15,799. From univariate sensitivity analysis it appears that the results are highly sensitive for the costs of treatment with captopril and the occurrence and prevention of clinical heart failure. Varying all estimates randomly between upper and lower limits - in 5000 simulations - an estimate of costs per life years gained of DFL 15,729 is made for 20 years of treatment, with 95% of all estimates between DFL 0 and DFL 50,000. On a national level undiscounted costs are expected to increase up to approximately DFL 42 million annually during the first 40 years after introduction of the preventive strategy.

* Based on Michel BC, Al MJ, Remme WJ, Kingma JH, Kragten JA, Nieuwenhuizen R van, Hout BA van. Economic aspects of treatment with captopril for patients with asymptomatic left ventricular dysfunction in the Netherlands. European Heart Journal. Accepted for publication.

Introduction

Heart failure is a relatively common disease in the Netherlands. With a prevalence of 13.5 per 1000 men and 8 per 1000 women between 35 and 85 years of age, it is estimated that heart failure is present in more than 80,000 persons.[1,2] In the near future this number will probably even increase due to the expected increase in the number of elderly where heart failure is relatively more prevalent.[1,3] Next to this demographic factor, a further reduction in early mortality after myocardial infarction is likely to cause an additional increase in the prevalence of heart failure.[1]
Due to the high prevalence of the disease in the Dutch population, successful strategies aimed at a reduction of the progression, the incidence or the prevalence of symptomatic heart failure will have large consequences.

Detection and treatment of risk factors - like hypertension, coronary artery disease, diabetes, obesity, serum cholesterol and smoking - are examples of strategies that may delay or even prevent the progression to symptomatic heart failure. Here changes in life style, screening of high risk individuals and optimal treatment of risk factors will probably have their effect. Additional benefits may be expected from new medical strategies. In this respect the introduction of ACE-inhibitors in the 1980s already had a major effect on the prevalence and costs related to heart failure. It has been estimated that treatment with ACE-inhibitors for patients with symptomatic heart failure improves survival and decreases costs by postponing the progression to the more advanced - and more costly - stages of the disease.[2] Recently, the SAVE trial and the SOLVD Prevention Trial have shown that treatment with ACE-inhibitors can also delay or prevent the occurrence of symptomatic heart failure.[4,5] Moreover, the SAVE trial, which investigated patients with a diminished left ventricular function after a myocardial infarction, also showed a significantly lower death rate by cardiovascular causes.[4]

In the Netherlands, preventive therapy with ACE-inhibitors is not common yet. In view of the economic aspects of this form of therapy several questions can be raised. Preventive treatment may be lifelong and it may involve large numbers of patients. Hence, costs may be considerable. On the other hand, there also is a potential to generate savings. Under the current treatment policy annually almost 429 million Dutch Guilders (about 200 million ECU) is spent on the treatment of symptomatic heart failure, which equals 1.1% of the total costs of Dutch health care.[6]

In this article the results of the SAVE trial are used to estimate the balance between costs and effects of preventive treatment with captopril, compared to the current strategy, for patients with asymptomatic left ventricular dysfunction after a myocardial infarction. Also the macro-economic consequences of an initiation of preventive treatment on a nation-wide level are estimated.

Methods

The SAVE trial evaluated the effects of a preventive treatment with captopril in two groups of about 1100 patients with asymptomatic left ventricular dysfunction after myocardial infarction. The follow-up of the trial was 42 ± 10 months, 82 percent of the patients were males and the average age of the patients was 60.[4] This trial showed that long-term administration of captopril is associated with an improvement in survival, a decrease in the incidence of symptomatic heart failure and a decrease in the number of reinfarctions.[7]

In the present study two base line scenarios were defined for the comparison of costs and effects between preventive treatment with captopril and the current strategy. The first scenario models the current treatment policy. In this scenario the probabilities concerning effects are obtained from the results reported for the placebo-arm of the SAVE trial. The second scenario models the situation in which captopril is subscribed. Here, the parameters concerning effects are taken from the results in the captopril-arm of the SAVE-trial.

To compare the costs and effects of these two scenarios a mathematical model was built. There were two reasons why this modelling technique was chosen.
First, data on costs were not made available in publications about the SAVE trial. Therefore, different sources of data had to be brought together and for such purpose a model is a suitable instrument.
The second reason is that preventive treatment can be seen as an investment for the future and that consequently, the balance between costs and effects needs to be assessed over a longer period than the trial. Again, a mathematical model is a suitable instrument to make such an extrapolation.

The model that was used is a so called Markov chain model, that incorporates data about the current strategy, the effects of captopril, the costs of treatment and

the epidemiology related to the patient population studied. With regard to the effectiveness of the therapy the model is mainly based on data from the SAVE trial.[4,8,9] With respect to the current treatment policy, the costs and the epidemiology, priority was given to data specific for the Dutch situation.

Two separate analyses were made. In the first analysis the results after four years were evaluated and compared to those of the SAVE trial using the cumulative event rates that were presented by Pfeffer et al.[4] In the second analysis the results were extrapolated to a period of 20 years.

THE MODEL

Figure 1 presents an outline of the model. The model evaluates the diagnostic and therapeutic events that occur in 100 patients with the characteristics of the patients in the SAVE trial during 4 (the first analysis) or 20 (the second analysis) consecutive years. For the first 4 years of treatment the model copies the SAVE results as closely as possible. After that, the model extrapolates the results with the assumption that the average effectiveness and treatment policy over the first four years will continue in the later years of therapy.

The model describes events and treatments for 100 patients who experienced a recent myocardial infarction but have no symptoms of heart failure. The layout of the model is identical for both scenarios with respect to the events that may occur and the treatments that are given. It is assumed that additional echocardiography (and sometimes radioventriculography) will have to be performed in patients who did not receive echocardiography as a routine follow-up of their infarction. The costs for this additional assessment are included in the captopril scenario. During the first year, patients with a-symptomatic heart failure experience normal follow-up of their index infarction, including analysis of cardiac ischaemia with exercise tests or thallium scans, angiography if one of these tests is positive, and cardiac intervention if the angiography indicates that this might be necessary, in both scenarios. Because the policy on the timing of interventions differs for separate hospitals and in separate periods, depending on the waiting list and on the urgency of the indication, no distinction is made as to whether the analysis for cardiac ischaemia and subsequent interventions were performed during the initial hospitalisation phase for the index infarction or during the follow-up phase, requiring readmission for the intervention. All surviving patients are assumed to

have a follow-up of their primary infarction consisting of, on average, 4 out-patient consultations, each including an electrocardiogram, medication and some basic laboratory tests. Some patients will also follow a cardiac rehabilitation programme. During subsequent years some patients will be re-evaluated for their cardiac function with an echocardiography and / or a radioventriculography. In some patients with stable angina pectoris, cardiac ischaemia will also be analysed and, if necessary, treated. The model includes unstable angina pectoris, re-infarction, cerebrovascular accident and symptomatic heart failure as the main cardiovascular complications. Every year a patient can have more than one complication. The events and treatments patients may encounter after re-infarction are shown in figure 2. The events and treatments for patients with unstable angina pectoris are identical, except for the thrombolysis.

After developing symptomatic heart failure, patients enter a separate part of the model. This part of the model was built earlier to estimate costs and effects of ACE-inhibitors for symptomatic heart failure, and includes the same cardiovascular complications as the main model.[2,10] Here patients are categorised by their condition according to the New York Heart Association classification. It is emphasised that patients in this part of the model who subsequently become asymptomatic do not re-enter the main model. As long as they have no symptoms they remain in New York Heart Association class I. Table 1 gives an overview of the data included in the model.

Figure 1 *Outline of the simulation model. CVA = cerebrovascular accident; RI = reinfaction; UAP = unstable angina pectoris; NYHA = classification by the New York Heart Association.*

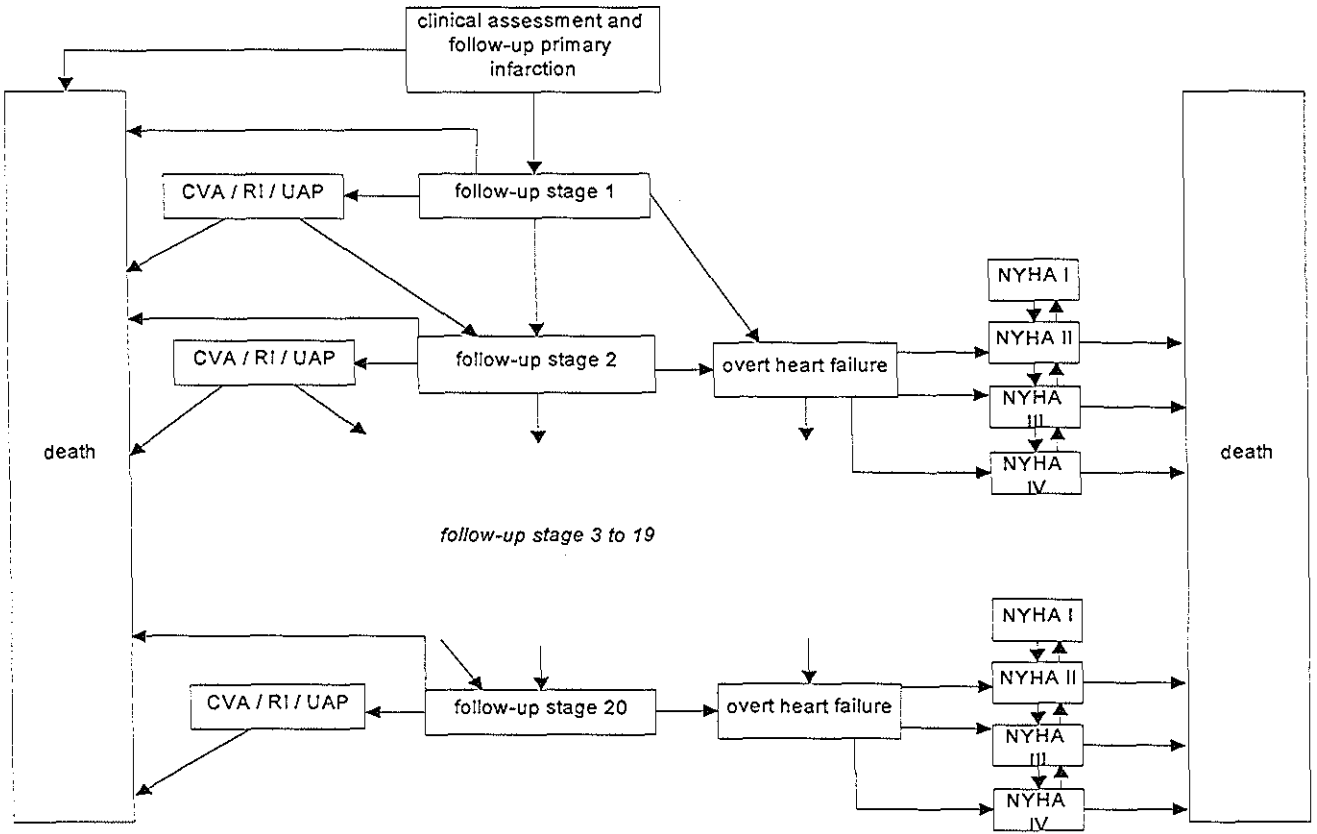


Figure 2 *Outline of the part of the model focusing on re-infarction*

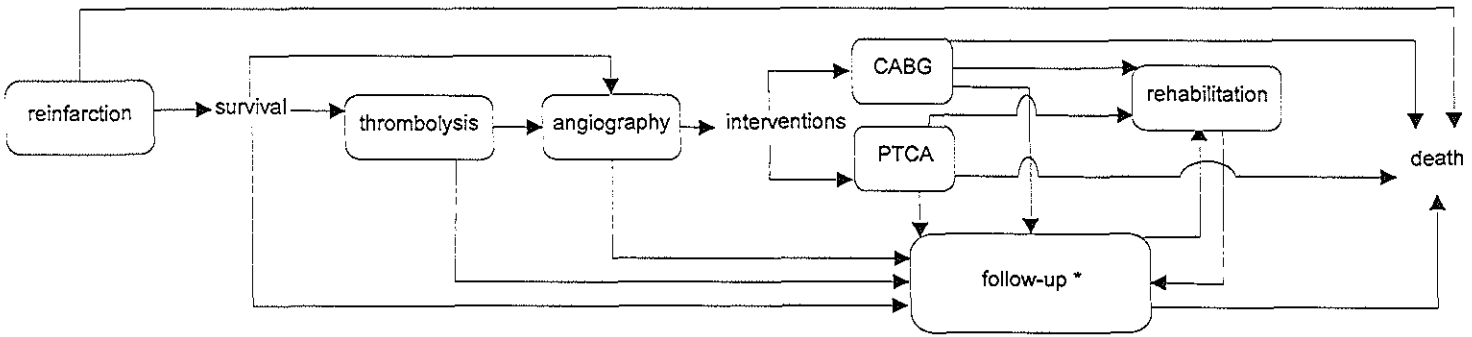


Table 1 *Minimum and maximum transition probabilities in the baseline scenario.*

	current strategy			preventive treatment capt.[a]		
	%	min	max	%	min	max
Echocardiography						
1st year	90	80	100	100		
next years	37.5	25	50			
Radioventriculography						
1st year	22.5	5	40			
next years	11.3	2.5	20			
Exercise tolerance test (ETT)						
1st year	90	80	100			
next years	40	20	60			
percentage of ETT with thallium	20	10	30			
Intervention						
1st year	10	5	15			
next years[b]	3	1	5			
percentage CABG/PTCA[c]	100	75	125			
Percentage angiography/ interventions[d]	225	150	300			
Patients' rehabilitation[11]						
after PTCA	25	20	30			
after CABG	50	45	55			
after uncomplicated infarction	33	25	40			
Percentage of patients with unstable angina who underwent subsequent angiography	55	40	70			
Reinfarction[e]						
in 1st year	8.5	6.9	10.1	7.5	6.0	9.0
in 2nd year	4.4	3.1	5.7	2.5	1.5	3.5
in 3rd year	3.5	2.3	4.7	2.5	1.5	3.5
in 4th year	3.1	1.7	4.5	2.0	0.9	3.1
in following years	3.7	1.4	5.9	2.3	0.6	4.1
Percentage of patients with a reinfarction treated with thrombolysis	32.9	22.9	42.9			
Cerebrovascular accident (CVA)	2	0.8	3.2			

	current strategy			preventive treatment capt.[a]		
	%	min	max	%	min	max
Symptomatic heart failure (CHF)						
in 1st year	22.0	19.6	24.4	16.0	13.9	18.2
in 2nd year	5.1	3.7	6.5	6.0	4.5	7.5
in 3rd year	6.1	4.6	7.7	2.8	1.7	3.9
in 4th year	6.5	4.5	8.5	6.5	4.5	8.5
in following years	5.9	3.1	8.8	5.1	2.4	7.8
inflow in NYHA III	30	25	35			
NYHA IV	20	15	25			
Death						
after coronary angiography	0.02	0.01	0.03			
after ETT	0.02	0.01	0.03			
after PTCA[12]	1	1	2.1			
after CABG[12]	3.6	1.1	3.6			
after unstable angina pectoris	11	9	13			
after reinfarction	22	15	35			
before diagnosis and therapy[f]	60	50	70			
after CVA in hospital	23.9	15	30			
CHF NYHA I	7.5	2.5	12.5			
CHF NYHA II	12.5	7.5	17.5			
CHF NYHA III	15	10	20			
CHF NYHA IV	40	25	45			

[a] *Preventive treatment with captopril. Only values that were different from those used in the "current strategy" are presented.*

[b] *In patients without unstable angina pectoris, re-infarction or symptomatic heart failure*

[c] *This means that in the baseline scenario the number of CABGs was equal to the number of PTCAs.*

[d] *This means that in the baseline scenario the number of angiographies was 2.25 times the number of interventions.*

[e] *For unstable angina pectoris the occurrence in both strategies is assumed to be identical to the occurrence of re-infarction in the preventive treatment strategy.*[7]

[f] *This means that it is assumed that in the baseline scenario 60% of the patients who die after a reinfarction will do so before costs are made for diagnosis and therapy.*

EFFECTS AND EPIDEMIOLOGICAL PARAMETERS

The two scenarios differ in terms of the probabilities concerning survival, the development of symptomatic heart failure, and the incidence of myocardial infarctions. These differences are based on the results of the SAVE trial.[4] The

probabilities related to the incidence of cerebrovascular accidents and unstable angina pectoris are similar in both scenarios.[7] In the 20 year extrapolation effects are expressed in terms of life years gained. In the 4 year analysis effects are also expressed in terms of additional survivors, because the period is rather short to give a reliable estimation in terms of life years gained.

With respect to the survival probabilities distinction is made between cardiovascular and non-cardiovascular causes of death. The cardiovascular death rates are based on the results of the SAVE trial.[4] The non-cardiovascular death rates, with a correction for cardiovascular mortality, are based on the survival tables of the Dutch Central Bureau of Statistics (CBS).[13,14] The time-dependent death rate due to a cerebrovascular accident is calculated according to Niessen et al.[15]

The probabilities that are relevant to the symptomatic heart-failure part of the model are taken from a previous study addressing the use of ACE-inhibitors in symptomatic heart failure. For that study the results of the SOLVD Treatment Trial were used.[16] To correct for the differences between trial results and daily practice it was assumed that 85% of the patients with symptomatic heart failure are treated with ACE-inhibitors and that only 60% of these patients benefit from this treatment according to the effectiveness seen in the SOLVD treatment trial.

COSTS

Costs were calculated from a societal point of view. Cost calculations were based on estimates of real costs, but limited to the medical costs of both treatments. Non-medical costs, direct and indirect (due to production losses), were not taken into account.

Information on the current diagnostic and therapeutic policy after myocardial infarction in the Netherlands was gathered by studying the literature and by a review of current practice in four non-academic hospitals, with between 424 and 1080 beds each (the average Dutch hospital has 422 beds). All cardiovascular events and all diagnostic or therapeutic actions were expressed in terms of units of resource use and multiplied with estimates of unit costs for the Netherlands. Costs of medication were based on officially published purchase prices for pharmacists.[18] Because it is estimated that 65% of the captopril used in the Netherlands is

imported against lower costs, the costs of treatment with captopril were adjusted for this.

Table 2 Cost-estimates included in the model

	Costs in DFL
Echocardiography[17]	ƒ 101.50
Radioventriculography[17]	ƒ 320.00
Treatment with captopril[18]	ƒ 1,846.68
Exercise test[11]	ƒ 121.00
Thalliumscan (tariff)	ƒ 923.10
Coronary angiography[19]	ƒ 2,825.00
PTCA[11]	ƒ 9,000.00
CABG[11]	ƒ 27,230.00
Rehabilitation[11,20]	ƒ 1,000.00
Follow-up after myocardial infarction	
1st year	ƒ 1,533.60
next years	ƒ 774.35
Unstable angina pectoris uncomplicated[11]	ƒ 4,608.00
Additional follow-up after unstable angina pectoris	ƒ 450.50
Reinfarction uncomplicated[11]	ƒ 9,729.00
Extra follow-up after reinfarction	
1st year	ƒ 380.00
next years	ƒ 759.00
Thrombolysis after reinfarction (r-TPA)[18]	ƒ 2,500.00
Cerebrovascular accident	
patient dies in hospital[21]	ƒ 8,908.00
patient survives[6,21]	ƒ 45,300.00
Symptomatic heart failure[2]	
NYHA I	ƒ 3,828.00 to ƒ 6,483.00
NYHA II	ƒ 7,818.00 to ƒ 9,313.00
NYHA III	ƒ 19,261.00 to ƒ 16,776.00
NYHA IV	ƒ 69,508.00 to ƒ 44,921.00

Table 2 presents an overview of the cost-estimates that are used in the model.

COST-EFFECTIVENESS

The cost-effectiveness of captopril treatment is defined as the additional costs that are made in the new scenario compared to the current scenario, divided by the number of additional survivors or life years that are gained (LYG) in the new scenario compared to the current scenario.

$$\frac{costs_{captopril} - costs_{current\ strategy}}{LYG_{captopril} - LYG_{current\ strategy}}$$

Here, this cost-effectiveness ratio is calculated in two ways. The cost effectiveness ratios after four and twenty years are estimated in terms of the costs per life year gained. This is calculated as the difference in costs made during 4 (20) years divided by the difference in life years gained in those 4 (20) years. When addressing the results after four years emphasis is placed on the costs per additional survivor after four years. This is calculated as the difference in terms of costs made during those four years divided by the difference in survival after four years. Both costs and effects are discounted by 5%, according to the guide-lines for cost-effectiveness analyses from a societal point of view.[22]

SENSITIVITY ANALYSIS

Two techniques are used to assess the uncertainties surrounding the baseline outcomes. The first technique, an univariate sensitivity analysis, addresses the sensitivity of the resulting cost-effectiveness ratio with respect to the various estimates by varying the estimates one by one. In the second technique, a multivariate sensitivity analysis, all estimates are varied simultaneously. For this multivariate sensitivity analysis upper and lower estimates were defined for all underlying estimates. The estimates concerning costs were surrounded with uncertainty margins of plus and minus 10 percent. The estimates concerning effectiveness were converted into 95% confidence intervals on the basis of the SAVE trial. Subsequently 5000 simulations were run, each leading to different cost-effectiveness ratios for the results after 4 respectively 20 years. By reporting the

distribution of costs, effects and cost-effectiveness ratios an impression is obtained of the reliability of the estimates.

CONSEQUENCES ON A NATIONAL LEVEL

The estimation of the consequences of the introduction of a preventive treatment policy on a national level is based on the intermediate population forecasts made in 1993 by the Dutch Central Bureau of Statistics.[23] It is assumed that the incidence of primary myocardial infarctions and reinfarctions together remains stable at 0.2% annually and that 7.25% of all infarctions are due to reinfarction.[11] No discount factor is used in the evaluation of the consequences on a national level.

Results

CONSEQUENCES ON A PATIENT LEVEL

If preventive treatment with captopril is introduced in a patient population, the model predicts effects that are comparable to the results of the SAVE trial. The model estimates a cumulative survival after 4 years of 74.8% in the current Dutch strategy and 79.2% after introduction of the new strategy, as compared to about 73.5% and 78.1% in the SAVE trial.[4] The model also estimates that the costs of medication will increase with DFL 4,937 per patient, whereas the costs of concomitant treatment will decrease with DFL 2,446 per patient. Consequently the total costs per patient are estimated at DFL 2,491 per patient after four years. Overall survival (after discounting) is estimated to increase from 61.5% to 65.1% after four years and so, the additional costs per additional survivor after four years are estimated at DFL 69,126. Additionally, it can be estimated that during the first four years, on average, 0.11 life years will be gained by preventive treatment. Consequently, the additional costs per life year gained during the first four years are estimated at DFL 22,887.

When extrapolating the results to 20 years, it is estimated that without captopril 9.4% of the patients with asymptomatic left ventricular dysfunction after a

myocardial infarction will still be alive. With captopril this is expected to improve to 12.8%. Additionally, it is estimated that, on average, 0.55 life years per patient will be gained. Finally, it is estimated that the total costs of treatment according to the current strategy will be DFL 59,419 per patient. With the preventive treatment policy used in the SAVE trial this amount is expected to increase to DFL 68,142. This results in an estimate of cost per life year gained of DFL 15,799.

SENSITIVITY ANALYSIS

The results estimated by the model are surrounded by several uncertainties. To evaluate those, all variables were changed one by one to ascertain the influence of each variable on the outcomes. This procedure shows that the estimation of the cost-effectiveness is extremely sensitive to the occurrence and prevention of clinical heart failure and the costs of treatment with captopril. For all other variables a change of 10% in the value of the variable results in a change in the 20 year cost-effectiveness ratio of less than DFL 1,200. In the baseline analysis it was assumed that 65% of the current consumption of captopril in the Netherlands is imported at a lower price. The argument can be made that this lower price of the import-drug is a better reflection of the price as it would be in a market of free enterprise, thereby representing societal costs more closely. If that argument is followed average additional costs after 4 years can be estimated at DFL 1,947 per patient, and the costs per additional survivor after 4 years at DFL 54,037. Concurrently, the additional costs per patient after 20 years are estimated at DFL 7,492 and the costs per life year gained at DFL 13,569.

In the baseline analysis it was also assumed that a high percentage (90%) of patients receiving standard care would have an echocardiography after their index infarction. If this percentage was to lower to, for example, 50% the cost-effectiveness ratio would be a little less favourable: 72,015 per additional survivor at four years and 15,961 per life year gained at 20 years.

In the 5000 simulations of the multivariate analysis the costs and effects vary according to the distributions shown in figure 3. The distributions of the cost-effectiveness ratio's are shown in figure 4. With respect to the four-year results the median cost-effectiveness ratio is DFL 69,651 per additional patient alive. Ninety-five percent of all estimates have a cost-effectiveness ratio between DFL 0 and DFL 240,000. With respect to the 20 year treatment the median costs per life year

gained are estimated at DFL 15,729 and here 95% of all estimates lay between DFL 0 and DFL 50,000.

Figure 2a Distribution of additional costs in 5000 simulations

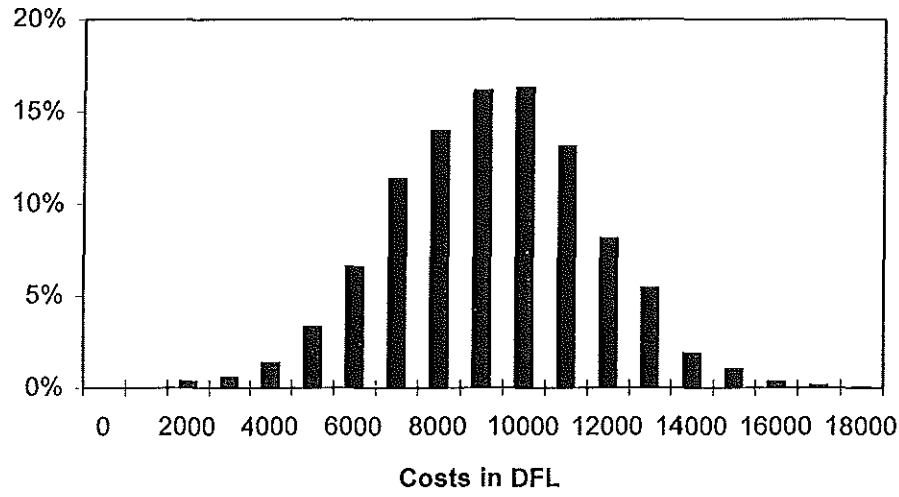


Figure 2b Distribution of additional effects (in life years gained) in 5000 simulations

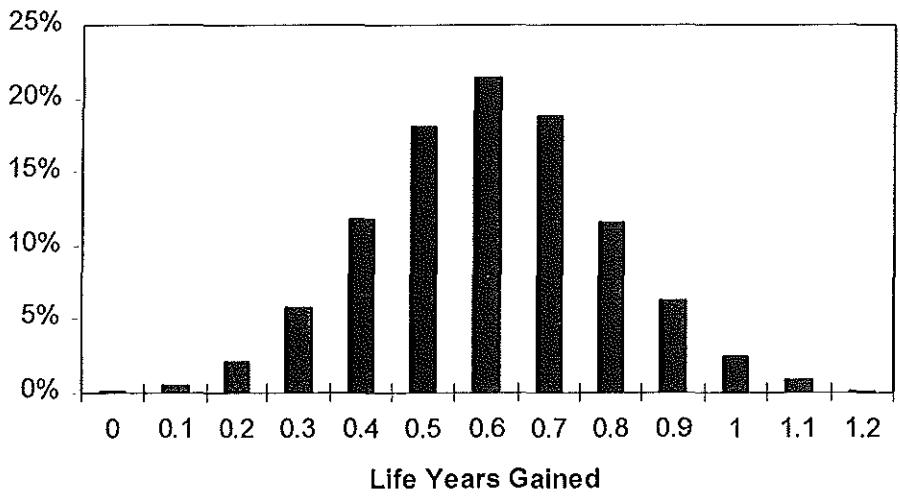


Figure 3a Distribution of 4 years cost-effectiveness ratio's in 5000 simulations

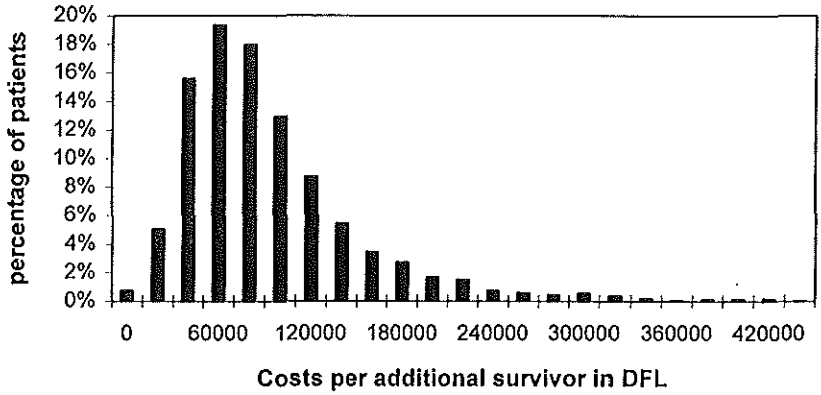
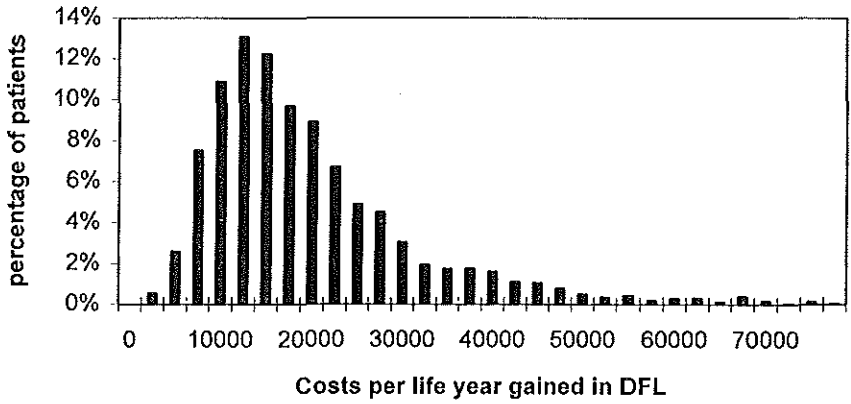


Figure 3b Distribution of 20 years cost-effectiveness ratio's in 5000 simulations



CONSEQUENCES ON A NATIONAL LEVEL

In the SAVE trial only 7% of the patients with a myocardial infarction qualified for preventive treatment with captopril. However, the inclusion criteria for the trial were very strict. For the current analysis it is estimated that 10% of the patients who are admitted into hospital for myocardial infarction will be eligible for preventive treatment. Figure 5 illustrates that costs will increase steadily during the

first years after introduction of the new policy. This is due to the larger number of people surviving as a result of the treatment. Costs are expected to remain increasing until a steady state is reached. If the strategy would have been introduced in 1994, this steady state could have been reached around the year 2035. In that year about 37,000 patients would be treated with captopril and each year 2,825 life years would be gained. On a national level the annual additional costs of preventive treatment would eventually amount to about DFL 42 million. If not 10%, but 20% or even 30% of the patients are eligible for preventive treatment this amount increases proportionately to DFL 84 million or DFL 125 million.

Figure 4a *Additional costs over the years by preventive treatment policy with captopril initiated in 1994 in the Netherlands*

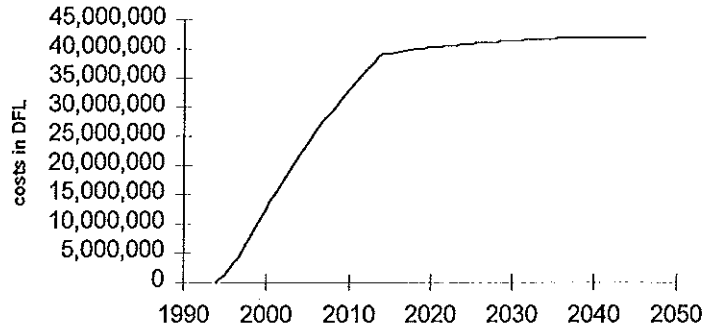
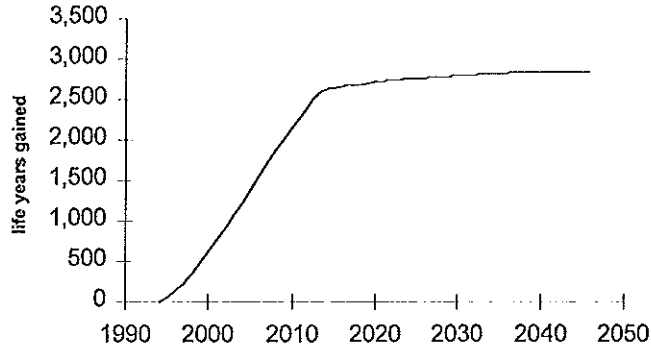


Figure 4b *Additional life years gained over the years by preventive treatment policy with captopril initiated in 1994 in the Netherlands*



Discussion

The current analysis shows that preventive treatment with captopril for patients with asymptomatic left ventricular dysfunction is likely to have positive effects. But costs will have to be made in order to gain these effects. Estimates based on the assumption that the results of the SAVE trial can be obtained in clinical practice point to a cost-effectiveness ratio of DFL 69,126 per additional survivor after 4 years of treatment. To give an estimation of the costs per life year gained the SAVE results were extrapolated to 20 years, resulting in a cost-effectiveness ratio of DFL 15,799 per life year gained. Here, several issues warrant comment.

First, the estimates in the model are surrounded by several uncertainties: cost data about the SAVE trial were not published, the results of the SAVE trial were extrapolated to the Dutch situation and a further extrapolation to a 20 year period was made to enable a reliable estimation of costs per life years gained. If account is taken of all confidence intervals surrounding the effects, treatment patterns and costs included into the model, an almost unlimited number of combinations of values for the separate items in the analysis can be made by computer simulation. In 5000 simulations the median cost-effectiveness ratios were compatible to those of the baseline-analyses: DFL 69,651 per additional survivor with a 4 year time-limit - with 95% of the estimates between DFL 0 and DFL 240,000 - and DFL 15,729 per life year gained with a 20 year time limit - with 95% of the estimates between DFL 0 and DFL 50,000. In the univariate sensitivity analysis it is shown that the occurrence of clinical heart failure and the level of prevention of this occurrence by captopril have a major influence on the cost-effectiveness ratio. Also, the price of treatment with captopril turns out to be important. The cost-effectiveness ratio will become more favourable if the price of captopril is lowered, e.g. after expiration of the patent. The dosage of captopril used in the SAVE trial was rather high: 79 % of the patients in the captopril group reached the target dosage of 150 mg per day. If preventive treatment with captopril is introduced with a lower standard dosage this may have a favourable effect on the cost-effectiveness ratio. However, such a lowering in dosage could also change the effects of treatment, thereby changing the cost-effectiveness ratio.

Secondly, the extrapolation of the results of the SAVE trial to a 20 year period is based on the assumption that the level of medical care will remain unchanged during this period. New technologies and treatment policies may have their

influence both on costs and effects during this period. The effect of this influence is unknown.

Thirdly, it should be noted that if preventive treatment is introduced on a wide scale the cost-effectiveness may be less favourable, because not all patients will be selected according to the strict inclusion criteria of the SAVE trial.[4]

Fourthly, the current analysis does not take effects on the quality of life and on indirect costs into account, because no data were available on these items. By reducing the number of patients with clinical heart failure, captopril, especially in the more severe stages, may have an important effect on the quality of life of the patient population involved. In the 20 year analysis the number of (undiscounted) life years spent in NYHA class IV is reduced by 7.7%, in NYHA class III by 9.2%, and in NYHA class II by 10.5%.

Finally, the current analysis is only valid for patients with asymptomatic left ventricular dysfunction after myocardial infarction, one of the major risk factors for clinical heart failure. Alternative preventive strategies may consist of treatment of all patients with acute myocardial infarction or of all patients with asymptomatic left ventricular dysfunction. Whereas the CONSENSUS-II trial did not show an improvement in survival after the use of ACE-inhibitors in an unselected patient population after myocardial infarction, the GISSI-3 and the ISIS-4 trial did show a decrease in mortality.[24,25,26] The ISIS-4 trial, which included more than 58,000 patients, showed that 1 month treatment with captopril, in a lower dose (maximum 100 mg per day), started within 24 hours of the onset of symptoms of suspected acute myocardial infarction may lead to a significant reduction in 5-week mortality. The benefit of 4.9 fewer deaths in 5 weeks (and 5.4 fewer deaths in 1 year) per 1000 patients treated with 1 month of captopril that was seen in the ISIS-4 trial, exceeds the benefit of 4.2 fewer deaths in 4 years per 1000 continuously treated patients seen in the SAVE trial. However, these two patient populations are not completely comparable. Patients in the ISIS-4 trial were unselected, whereas patients in the SAVE trial had a proven left ventricular dysfunction. Further research will have to show which dosage, moment of initiation of therapy and duration of treatment is most cost-effective in patients with a recent myocardial infarction. In contrast to the previously mentioned trials, the SOLVD Prevention trial concentrated on the effect of ACE-inhibitors in patients with asymptomatic left ventricular dysfunction who did not experience a recent myocardial infarction. This trial showed an effect on the occurrence of symptomatic heart failure but no

significant effect on the cardiovascular death rate.[5] Probably, the screening that is necessary to find symptomless patients without overt risk-factors in the general population, like the patients included in the SOLVD Prevention trial, will make the cost-effectiveness of this treatment policy less favourable than the treatment of patients who are at-risk for a decrease in left ventricular function because they recently experienced a myocardial infarction.

In conclusion it can be stated that, if only the effects are taken into account, all signs seem to be in favour of the introduction of preventive treatment of patients with asymptomatic ventricular dysfunction after an acute myocardial infarction in the Netherlands. In addition to the health care aspect, preventing or delaying the onset of heart failure, the strategy shows positive short-term and long-term effects on survival. If the positive effects are related to costs the picture is less favourable. But, even though the cost-effectiveness of only few preventive treatments was analysed for the Netherlands, most of them were shown to be less cost-effective. For the Netherlands the costs of screening for cervix carcinoma are estimated at DFL 24,000 per life year gained, and of cholesterol lowering therapy for men with cholesterol level above 8 mmol/l at about DFL 30,000 per life year gained.[27,28] However, the costs of screening for breast cancer are lower than the amount estimated here: DFL 7,650 per life year gained.[29]

Health care authorities will have to decide whether the current cost-effectiveness ratio is favourable enough to warrant introduction of preventive treatment without additional trials focusing on the gathering of data on costs and effects specific for the Dutch situation and whether they can afford the additional costs involved in the new strategy that could amount to up to DFL 42 million per year: 10.3 percent of the amount spent on heart failure in 1988.

References

1 Bonneux L, Barendregt JJ, Meeter K, Bonsel GJ, Maas PJ van der. Estimating clinical morbidity due to ischemic heart disease and congestive heart failure: the future rise of heart failure. Am J Pub Health 1994; 84(1): 20-28

2 Hout BA van, Wielink G, Bonsel GJ, Rutten FFH. Effects of ACE inhibitors on heart failure in the Netherlands. Pharmacoeconomics 1993; 3(5): 387-97

3 Reitsma JB, Mosterd A, Koster RW, Capelle FJL van, Grobbee DE, Tijssen JGP. Increase of the number of hospital admissions for heart failure in Dutch hospitals between 1980 and 1992 (in Dutch). Ned Tijdschr Geneeskd 1994; 138(17): 866-71

4 Pfeffer MA, Braunwald E, Moyé LA, et al. Effect of captopril on mortality and morbidity in patients with left ventricular dysfunction after myocardial infarction. New Eng J Med 1992; 327(10): 669-77

5 The SOLVD investigators. Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions. New Engl J Med 1992; 327(10): 685-91

6 Koopmanschap MA, Roijen L, Bonneux L. Costs of illness in the Netherlands (in Dutch). Rotterdam: instituut Maatschappelijke Gezondheidszorg / instituut voor Medische Technology Assessment, 1991

7 Rutherford JD, Pfeffer MA, Moyé LA, et al. Effects of captopril on ischemic events after myocardial infarction. Results of the Survival and Ventricular Enlargement trial. Circulation 1994; 90(4): 1731-1738

8 Rouleau JL, Moyé LA, Pfeffer MA, et al. A comparison of management patterns after acute myocardial infarction in Canada and the United States. New England J Med 1993; 2328(110: 779-84

9 Arthur D Little. Bristol-Myers Squibb Economic analysis of captopril in slowing the rate of progression to congestive heart failure. Draft final report. November 1993.

10 Hout BA van, Wielink G, Bonsel GJ, Rutten FFH. Heart failure and ACE inhibitors. A cost effectiveness analysis. Rotterdam: institute for Medical Technology Assessment: 1992

11 Boas GM. Scenario analysis of the economic aspects of coronary heart diseases (in Dutch). Thesis, University of Limburg, 1994.

12 Breeman A, Serruys PW, Brand MJBM van den, Deckers JW, Herwerden LA van, Roelandt JRTC. Complications shortly after percutaneous transluminal angioplasty and after coronary surgery in 183 comparable patients with multiple vessel coronary disease (in Dutch). Ned Tijdschr Geneeskd 1994; 138(12): 1074-80

13 Netherlands Central Bureau of Statistics. Monthly bulletin of health statistics. February 1994. Voorburg/Heerlen: Netherlands Central Bureau of Statistics, 1994.

14 Netherlands Central Bureau of Statistics. Monthly bulletin of population. May 1993. The Hague: SDU/publishers, Netherlands Central Bureau of Statistics, 1993.

15 Niessen LW, Barendregt JJ, Bonneux L, Koudstaal PJ for the Technology Assessment Methods Project Team. Stroke trends in an aging population. Stroke 1993; 24(7): 931-9

16 The SOLVD-investigators. Effect of Enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. NEJM 1991; 325(5): 293-302

17 Hout BA van. Heart transplantations: costs, effects and prognosis (in Dutch). Thesis, Erasmus University Rotterdam, 1990.

18 Netherlands Health Insurance Executive Board. Pharmacotherapeutical Compass 1994. Amstelveen: Netherlands Health Insurance Executive Board, 1994.

19 Simoons ML, Vos J, Martens LL. Cost-utility analysis of thrombolytic therapy. European Heart Journal 1991; 12: 694-99

20 Gosselink KJM. Heart revalidation - a study into the costs - (in Dutch). Essay, University of Limburg, 1989.

21 Versloot GR, Meulen JHP van der, Limburg M, Hout BA van, Habbema JDF. Estimation of costs of hospital stay for stroke patients in the Netherlands. Submitted 1994.

22 Rutten FFH, Ineveld BM van, Ommen R van, Hout BA van, Huijsman R. Cost calculations in health care research. Practical guide-lines (in Dutch). Rijswijk: Steering Committee Future Scenarios in Health Care, 1993.

23 Netherlands Central Bureau of Statistics. Monthly bulletin of population. December 1993. The Hague: SDU/publishers, Netherlands Central Bureau of Statistics, 1993.

24 Gruppo Italiano per lo Studio della Sopravvivenza nell'Infarto Miocardico. GISSI-3: effects of lisinopril and transdermal glyceryl trinitrate singly and together on 6-week mortality and ventricular function after acute myocardial infarction. Lancet 1994; 343: 1115-22

25 Swedberg K, Held P, Kjeksjus J, et al. Effect of the early administration of enalapril on mortality in patients with acute myocardial infarction. Results of the Cooperative New Scandinavian Enalapril Survival study II (CONSENSUS II). N Eng J Med 1992; 327: 678-84

26 ISIS-4 Collaborative Group. ISIS-4: a randomized factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58050 patients with suspected acute myocardial infarction. Lancet 1995; 345: 669-685

27 Ballegooijen M van, Habbema JDF, Oortmarssen GJ van, Koopmanschap MA, Lubbe JThN, Agt HME van. Preventive pap-smear: balancing costs, risks and benefits. Br J Cancer 1992; 65: 930-3

28 Martens LL, Rutten FFH, Erkelens DW, Ascoop CAPL. Cost-effectiveness of cholesterol lowering therapy in the Netherlands: simvastatin versus cholestyramine. Am J Med 1989; 87: 54S-58S

29 Koning HJ de, Ineveld BM van, Oortmarssen GJ van, et al. Breast cancer screening and cost-effectiveness; policy alternatives, quality of life considerations and the possible impact of uncertain factors. Int J Cancer 1991; 49: 531-7

PLANNING OF NEONATAL ECMO FACILITIES*

Summary

This chapter illustrates the use of a model, constructed in an universally available spreadsheet programme, in the planning of the number of neonatal ECMO facilities. For any number of patients presented for treatment in a region annually and any number of ECMO facilities the number of patients that will have to be referred to facilities abroad is estimated in a microsimulation model. The input of the model consists of the number of ECMO facilities, the number of patients presented annually, the duration of treatment and the date that patients are presented for ECMO treatment. To illustrate the method, the Dutch situation in 1992 (29 patients treated in 3 facilities) was described and compared to several future scenarios. In the main future scenario it is assumed that the number of patients will increase to 56 patients per year. The model shows that no additional ECMO facilities are necessary in the Netherlands, if between two and three referrals to centres abroad are acceptable and feasible. In that situation it is expected that on 21 occasions two patients will be treated simultaneously, for a total of 79 days. On 10 occasions all three facilities will be occupied at the same time, for 20 days in total. During 201 days at least one of the facilities will be occupied. Although the results of the simulation are presented for the Netherlands, they are generally applicable to any region in any country with similar characteristics with respect to neonatal ECMO. For other regions the method is applicable, but the results may change.

Introduction

Extracorporeal membrane oxygenation (ECMO) is one of the new technologies in neonatal intensive care. It is potentially life saving for neonates with reversible

* Based on: Michel BC, Staveren RJE van, Geven WB, Hout BA van. Planning of neonatal ECMO facilities. Submitted.

respiratory diseases due to various causes. By temporarily taking over the oxygenation it can give diseased lungs opportunity to recover and it can avert the pulmonary complications that frequently occur in prolonged high pressure ventilation. However, ECMO itself may induce complications, most notably bleedings due to systemic heparinisation.[1,2] Therefore, treatment is limited to those who are thought to have a high risk of death with conservative treatment and a low risk of complications with ECMO.

The introduction of ECMO was realised differently in various countries. In the United States ECMO already is an accepted form of therapy for the above mentioned patient population.[3,4] In Europe ECMO was introduced more slowly. Since 1987 ECMO centres have been founded in Germany, France and Sweden.[5,6] In the United Kingdom ECMO was introduced recently in the context of a randomised controlled trial, the results of which are not available yet.[7] In the Netherlands the attitude towards neonatal ECMO has also been reserved. At the moment there are three centres, all waiting for a definite decision on their status by the health care authorities.

Once health care authorities approve of ECMO they also have to decide on the number of ECMO facilities that are necessary to provide adequate care. Additionally, in a region where ECMO already is an established therapy it may be necessary to adjust the number of ECMO facilities if changes in the indications for this treatment occur. It is to be expected that the required number of facilities in a region will depend on the number of patients who fulfil the indications for ECMO, the average duration of ECMO treatment and on the occurrence of simultaneous patient presentation. Therefore, for any given number of ECMO facilities, momentary over- or undercapacity may occur. Since ECMO is a potentially life-saving treatment, patients who can not be treated at moments of undercapacity will either have to be treated conservatively - with almost certainly a lower survival probability - or be referred to ECMO centres elsewhere. Whether or not such referrals are feasible and desirable are questions that will not be addressed in this article. Here, we concentrate on an approach to estimate the consequences of the establishment of various numbers of ECMO facilities in a certain region. To illustrate this approach the situation in the Netherlands is analysed. But, although the results of this estimation are presented for the Netherlands, they are generally applicable to any region in any country with the same number of patients requiring ECMO treatment, similar indications and the same distribution of ECMO treatment durations. For other regions the method is applicable, but the results may change.

Patients and methods

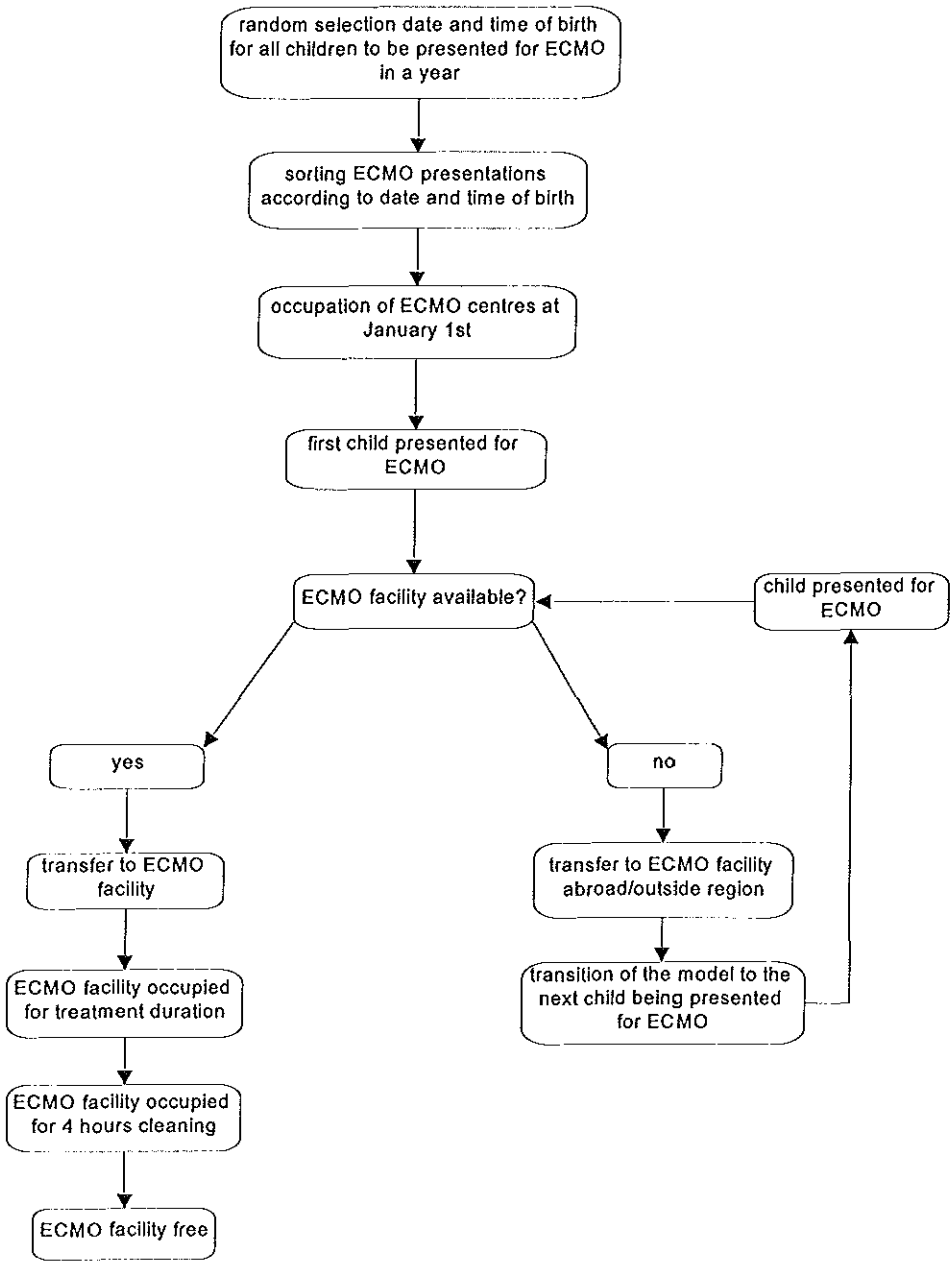
The number of ECMO facilities required to treat various numbers of patients was estimated by simulating ECMO treatments for 750 separate years and 1,361,250 individual patients in a discrete event micro-simulation model.(figure 1) The microsimulation model has four variables as input: the number of ECMO facilities, the number of patients presented annually, the duration of treatment and the date patients are presented for ECMO treatment. The model was developed in an universally available spreadsheet programme (Quattro Pro for Windows version 5.00, Borland International, Inc, 1993). First a baseline analysis was performed, describing the situation in the Netherlands in 1992. Then, several alternative scenarios were formulated to study the effects of future events on the number of ECMO-facilities required. Finally a sensitivity analysis for the main variable, duration of treatment, was performed.

THE BASELINE ANALYSIS

The baseline analysis was based on real data for the Netherlands. In 1992 three ECMO centres were functioning in the Netherlands. Two centres, in Rotterdam and Nijmegen, included patients strictly according to a common research protocol, limiting treatment to neonatal patients with an $AaDO_2$ above 600 mm Hg for 8 hours, or an Oxygenation Index above 40 in 3 to 5 consecutive blood gas samples within 5 hours. Patients with serious heart disease, causing an important right to left shunting, were excluded.

The third programme, in Maastricht, was not bound to an official protocol. For this centre only data were included on the 5 patients that would have fulfilled the entry criteria of the research protocol used by the other two centres. Data on all 53 patients, treated between January 1991 and July 1993, were included in the model. The main characteristics of these patients are described in table 1.

Figure 1 Outline of the simulation model



The mean treatment duration for these patients was 6 days with a minimum of 0.6 days and a maximum of 15 days. For the model the actual distribution of these times was used to simulate the duration of treatment for each patient. Besides, 4 hours were added to the duration of treatment, representing the minimum period necessary for cleaning and priming the equipment. Patients were said to be treated simultaneously if any part of their treatment period, including the cleaning period, overlapped with the same period of any other patient.

Table 1 Characteristics of the 53 patients treated between January 1991 and July 1993.

Characteristics	Patients / values
Diagnosis[a]	
- congenital diaphragmatic hernia	16 (30.2%)
- meconium aspiration syndrome	22 (41.5%)
- other	15 (28.3%)
Gestational age in weeks[b]	39.4 ± 1.8
Birth weight[b]	3325 ± 523
Hours on ECMO[b]	144 ± 73
Hours between birth and ECMO[b]	45 ± 55

[a] number of patients (%) ; [b] mean ± standard deviation

For each patient the model simulated the date of birth according to the monthly distribution of births in the Netherlands in 1992.[8] Here the assumption was made that ECMO indications themselves show no separate seasonal trend. As a matter of convenience the date of birth was used as an approximation of the day of presentation for treatment, because patients were treated with ECMO very soon after birth, within on average 1.9 days (0 to 15 days).

ALTERNATIVE SCENARIOS

Because ECMO was a new medical treatment facility in the Netherlands, referral policy for this treatment was probably still suboptimal in 1992. Also, it can be expected that inclusion criteria and contra-indications will loosen up in time. Therefore the number of treatments per year will probably rise. If the 29 treatments in 1992 are related to the 196,734 life born babies that year,[8] one ECMO treatment

Table 2 Results of the simulation model predicted for 1992 compared to the observed events in 1992, and several alternative scenarios. min = minimum; max = maximum; obs = observed events in 1992; NA = not available

	1992 situation 29 patients 3 facilities				alternative scenarios		
					56 patients 3 facilities	56 patients 4 facilities	56 patients 5 facilities
	average	min	max	obs	average	average	average
Patients referred for treatment abroad	0.3	0	2	NA	2.7	0.8	0.1
Occasions patients treated simultaneously (after referral):							
2 simultaneous	8.6	3	14	10	21.3	20.9	20.5
3 simultaneous	2.0	0	5	0	9.5	9.4	9.6
4 simultaneous						2.8	3.0
5 simultaneous							0.7
Days at least 1 patient is treated	133.9	84	173	168	200.9	199.3	203.5
Days patients are treated simultaneously							
2 simultaneous	32.3	11	65	32	78.8	79.2	75.9
3 simultaneous	4.1	0	19	0	20.4	20.8	21.3
4 simultaneous						4.8	4.5
5 simultaneous							0.8

in 6,784 life born babies was performed. Estimates based on the ELSO registry tend to a higher rate of ECMO treatments: one in 3,859 life born babies in 1988.[4] Investigators in Georgia, Michigan, and Oregon/Washington estimated the need for ECMO in the same magnitude: one in 3,717, one in 3,431 and one in 3,521 births respectively.[4,9] The need for ECMO in the UK is estimated at approximately 1 in 5,000 life born babies.[10]

If these estimates apply to the Dutch situation - after the treatment has become current practice - the number of treatments will rise to 39 or maybe even to 56 per year and of course the number of ECMO-facilities will have to be adapted to this new situation. In the alternative scenarios account is taken of these changes.

Results

THE BASELINE ANALYSIS

In the baseline analysis 29 patients per year are treated in 3 facilities, as was the situation in the Netherlands in 1992. The results from the simulations are summarised in the first three columns of Table 2. The estimates, based on the microsimulation model, differ slightly from the observed figures, which are shown in the fourth column, but all observed figures fall within the minimum-maximum range of the results estimated by the model. However, caution should be taken in comparing the estimates with the observed figures, because no record was kept of the referrals to centres abroad.

ALTERNATIVE SCENARIOS

The alternative scenarios simulate the effects of a change in the number of neonatal patients annually presenting for treatment or a change in the number of facilities, or both. It is assumed that there is no change in the average duration of treatment or in the distribution of treatment durations. As may be expected, the number of occasions in which patients require treatment at the same time increases with the number of patients presented per year.

Figure 2 The number of occasions 2 to 6 patients will be treated at the same time related to the number of patients requiring treatment annually in a situation with an unlimited number of ECMO facilities. (no. = number; yr = year)

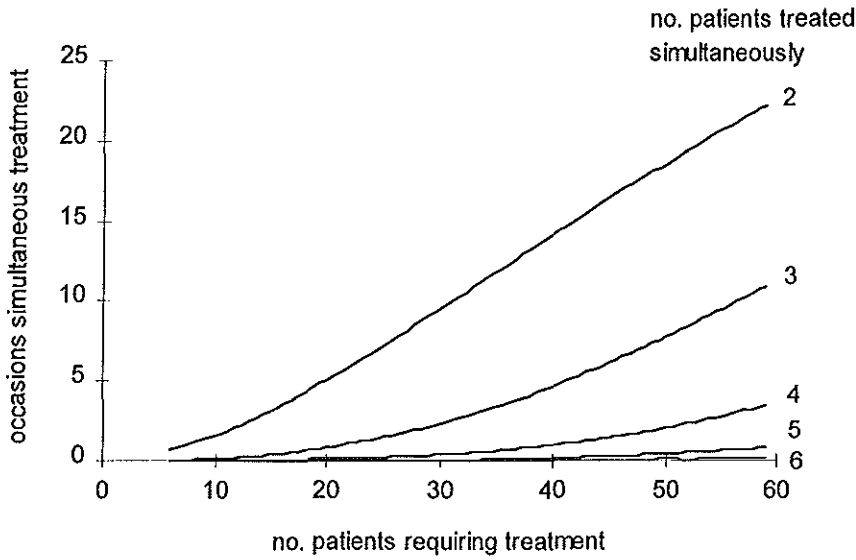
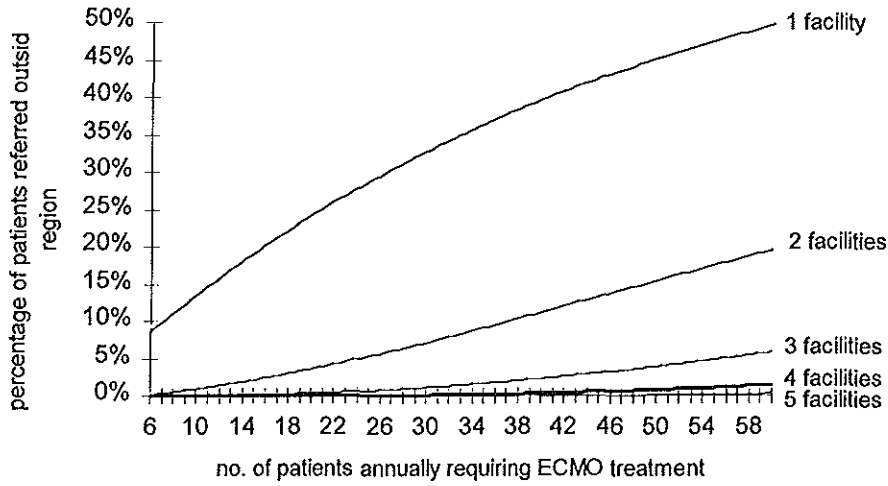


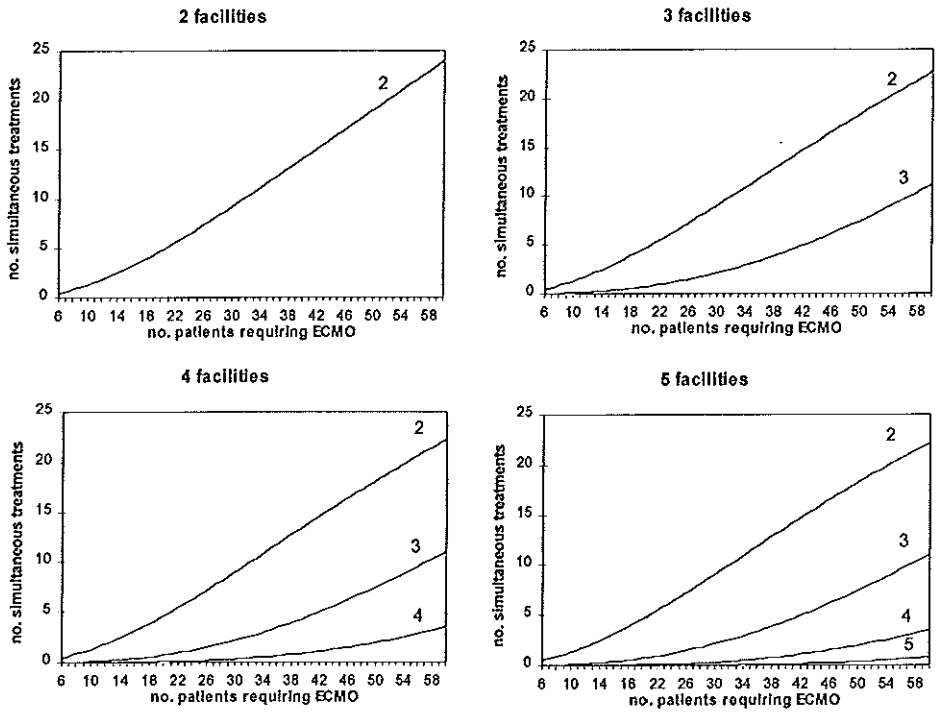
Figure 2 shows that, even if 60 patients would fulfil ECMO criteria annually, there would hardly ever be an occasion on which 6 patients would need treatment at the same time. However, the assumption that an unlimited number of facilities will be available does not seem realistic. It is highly probable that the number of treatment facilities will be limited, and that consequently some patients will have to be referred to facilities outside the region (or be treated conservatively). Depending on the number of facilities and the number of patients annually requiring treatment, it is estimated that between 0 and 30 patients will have to be referred to facilities outside the region.(figure 3) If only one facility is established, sometimes as many as 50% of the patients presented to that facility will have to be referred. In figure 4 the number of occasions that patients will be treated simultaneously is shown for 2 to 60 patients requiring treatment annually and for 2 to 5 facilities.

Figure 3 *The percentage of patients that will have to be referred to ECMO centres outside the region, if 1 to 5 facilities are available in the region and 1 to 60 patients require treatment annually. (perc. = percentage; no. = number; yr = year)*



Facing an increase to 56 patients annually, Dutch health care planners can decide to leave the number of facilities unchanged or expand to 4 or 5 facilities. Table 2 shows the consequences of these scenarios. An expansion from 3 to 4 facilities will diminish the number of referrals from 2.7 to 0.8 patients on average. It is expected that all 4 facilities will be occupied at the same time during 4.8 days per year. A further expansion to 5 facilities decreases the number of referrals even more, to only 0.1 patients. The disadvantage of such a policy is that the fifth facility will be occupied for only 0.8 days per year. Of course in reality the occupation of facilities will be more evenly distributed, due to the geographical spread of the facilities. If the presentation of neonates is spread evenly over the facilities, the 3 facilities in the baseline analysis will on average be occupied for 45 days per year.

Figure 4 The number of occasions 2 to 5 patients will be treated simultaneously, if 2 to 5 facilities are available, related to the number of patients annually requiring treatment. Patients that are referred to centres outside the region are included in the number of patients annually requiring treatment but excluded from the patients treated simultaneously. (no. = number; yr = year)



SENSITIVITY ANALYSIS

If the average duration of ECMO treatment is increased by one (16.7%) or two days (33.3%), while its distribution remains unchanged, the average number of referrals to centres outside the region in the 1992 situation increases to on average 0.4 patients per year, an additional 33%, or 0.5 patients, an additional 67%, respectively.(table 3) However, all changes in the average referrals found in the sensitivity analysis fall within the range of referrals presented in table 2.

Table 3 Results of the sensitivity analysis for the increase in the duration of treatment in the 1992 situation, and several alternative scenarios

	1992 situation		alternative scenarios					
	29 patients 3 facilities		56 patients 3 facilities		56 patients 4 facilities		56 patients 5 facilities	
Increase in the duration of treatment in days	+1[a]	+2[b]	+1	+2	+1	+2	+1	+2
Patients referred for treatment elsewhere	0.4	0.5	4.0	5.7	1.1	0.9	0.2	0.4
Occasions patients treated simultaneously (after referral):								
2 simultaneous	9.0	10.4	20.4	20.2	20.7	20.6	20.9	20.3
3 simultaneous	2.5	3.3	11.7	13.9	11.3	12.5	11.5	12.4
4 simultaneous					4.0	4.9	3.6	5.0
5 simultaneous							1.0	1.2
Days at least 1 patient is treated	150.2	166.6	219.8	238.7	224.7	248.7	227.1	241.7
Days patients are treated simultaneously								
2 simultaneous	37.0	48.3	89.3	97.4	86.2	92.6	88.9	93.4
3 simultaneous	5.9	8.6	28.4	37.0	28.8	33.6	27.9	37.6
4 simultaneous					7.7	8.6	6.8	9.9
5 simultaneous							1.1	2.0

[a] *+1 = an increase of 1 day;* [b] *+2 = an increase of two days*

DISCUSSION

Although the planning of all health care facilities is recommendable, the planning of ECMO facilities is particularly important since the patients involved can not be placed on a waiting list. Patients who can not be treated because of a shortage in facilities in the region will have to be treated conservatively or be referred to other regions. Planning is necessary not only when ECMO is introduced in a region but also when changes in the number of patients requiring ECMO occur. Such changes, both increases and decreases, could occur through changes in the indications after publication of the results of the randomised trials that are currently in progress, through changes in the alternative treatments and through improvements in the technique.[7,11,12] An increase in the annual number of patients may lead to higher levels of occupation of the facilities, causing an increase in referrals to facilities outside the region. If the number of ECMO facilities is subsequently increased the number of patients for whom facilities have to be found outside the area decreases. But, if too many facilities are established, money will be spent on facilities that will be left unoccupied for most of the year. A decrease in the number of patients will lead to structural overcapacity. Such overcapacity may loosen up criteria for ECMO treatment, causing an increased use by patients who do not need ECMO to survive, which in turn decreases the cost-effectiveness of the technology. Careful planning of the number of facilities may prevent these effects.

Planning problems in which random processes play a major role, like the presentation of ECMO patients, can be partially solved by microsimulation models. Of course models are abstract representations of reality, and, as stated by Feest and Harrison, they are only as good as the data fed into them.[13] But, although they have their limitations, they provide information that can be used in the planning of health care facilities. Most of the planning in health care with the aid of microsimulation was done for renal services and hospital bed requirements.[14-19] Unfortunately, the results were mainly published in journals specialised in operational research, thereby staying out of focus for most health care workers.[16,17,18,20]

The microsimulation model for the planning of ECMO facilities shows that in the Netherlands - if there is no change in the average duration of treatment - an increase to 56 patients a year can take place without an increase in ECMO

facilities, if between two and three referrals to centres elsewhere are accepted and feasible. The acceptability or feasibility of these referrals turn out to be a crucial point in the planning, because it is unknown who is to decide on these issues. If these referrals are not accepted or not feasible, an increase in facilities has to be balanced against an investment in facilities which will not be occupied frequently. The model also shows that even after an increase to 56 patients per year the marginal contribution of a fifth facility is very small. Due to the geographical spread of facilities and the patients the occupation may increase, but this increase will be at the expense of the occupation of the other 4 facilities.

In countries with large differences in the geographic distribution of patients requiring ECMO treatment, due to transport problems, geographic remote areas or non-uniformity of the patient distribution, the model can not be applied to estimate the number of ECMO facilities required on a national level, but it can be used to estimate the need for ECMO facilities in smaller sub-regions with a homogenic patient distribution. For the Netherlands, the geographical distributions of facilities and patients are not too important for planning purposes, because it is a small country and newborn babies are relatively easy transported in ambulances and helicopters. In fact, specialised neonatal care in the Netherlands is organised on a national rather than on a regional basis, with a central co-ordinating system for the occupation of neonatal facilities.

The current study shows that it is possible to plan the number of ECMO facilities required in a region with relative simple tools. Such planning may help ECMO centres in their decision whether or not to invest in extra facilities in a changing field. The study also shows that the acceptability and feasibility of referrals to centres outside the region - or abroad - is an important issue which European health care planners will have to consider.

References

1 Sell LL, Cullen ML, Whittlesey GC, Yedlin ST, Philippart AI, Bedard MP, Klein MD. Hemorrhagic complications during extracorporeal membrane oxygenation: prevention and treatment. J-Pediatr-Surg 1986; 21(12): 1087-1091

2 Trento A, Griffith BP, Hardesty RL . Extracorporeal membrane oxygenation experience at the University of Pittsburgh. Ann-Thorac-Surg 1986; 42(1): 56-59

3 Klein MD. Neonatal ECMO. ASAIO-Trans 1988; 34(1): 39-42

4 Schumacher RE. Extracorporeal membrane oxygenation. Will this therapy continue to be as efficacious in the future? Update on Neonatology 1993; 40(5): 1005-1022
5 Chevalier JY, Durandy Y, Batisse A, Mathe JC, Costil J. Preliminary report: extracorporeal lung support for neonatal acute respiratory failure. Lancet 1990; 335(8702): 1364-1366
6 Frenckner B, Ehren H, Palmer K. Extracorporeal membrane oxygenation in a European center-clinical experience. Eur J Pediatr Surg 1991; 1(1): 15-20
7 Field DJ, Firmin RK. Extracorporeal membrane oxygenation (letter). Lancet 1991; 337: 51-52
8 Durch Central Bureau of Statistics. Monthly bulletin of health statistics. February 1993. Voorburg/Heerlen: Dutch Central Bureau of Statistics, 1993
9 Southgate WM, Howell CG, Kanto WP Jr. Need for and impact on neonatal mortality of extracorporeal membrane oxygenation in infants of greater than 2500-gram birth weight. Pediatrics 1990; 86(1): 71-74
10 Greenough A, Emery E. ECMO and outcome of mechanical ventilation in infants of birthweight over 2 kg. Lancet 1990; 336(8717): 760
11 Donn SM. Alternatives to ECMO. Arch Dis Child 1994; 70: F81-F83
12 Kanto WP. A decade of experience with neonatal extracorporeal membrane oxygenation. J Pediatrics 1994; 124(3): 335-347
13 Feest TG, Harrison P. Simulation model for planning renal services (letter). BMJ 1992; 305:1018
14 Bolger PG, Davies R. Simulation model for planning renal services in a district health authority. BMJ 1992; 305: 605-608
15 Coats RB, Parkin A. Computer models in the social sciences. Bath: Edward Arnold, 1977
16 Davies R, Davies T. Using simulation to plan health services resources: discussion paper. J Roy Soc Med 1986; 79: 154-157
17 Harris RA. Hospital bed requirements planning. European Journal of Operational Research 1985; 25: 121-126
18 Romanin-Jacur G, Facchin P. Optimal planning of a pediatric semi-intensive care unit via simulation. European Journal of Operational Research 1987; 29: 192-198
19 Dumas MB. Hospital bed utilization: an implemented simulation approach to adjusting and maintaining appropriate levels. Health services research 1985; 20(1): 43-61
20 Davies H, Davies R. A simulation model for planning services for renal patients in Europe. J Opl Res Soc 1987; 38(8): 693-700

DISCUSSION AND CONCLUSIONS

In the previous chapters several practical applications of mathematical-statistical modelling techniques in Medical Technology Assessment were presented.

The use of a model in the *preparatory phase* of a RCT was illustrated with three examples. In the first example the whole study was modelled in order to find the optimal design of the RCT, in the second example a diagnostic tool was developed with modelling techniques. This tool was subsequently used in the third example, in which the most cost-effective diagnostic strategy to be included in a RCT was selected from a large amount of possible strategies.

In the first example a model was built to study the feasibility of the introduction of helicopters with medical crews for seriously injured accident victims in the Netherlands (chapter 1).[1] The introduction of these helicopter "trauma teams" was surrounded by a (rather polarised) discussion focusing on the costs and effects of the helicopter service itself. However, the formal modelling of the problem and the subsequent scenario analysis showed that not so much the uncertainty surrounding the costs of the helicopter service, but more so the uncertainty surrounding the gain in survival from the service and the costs of hospital treatment for seriously wounded trauma victims play a decisive role in determining the cost-effectiveness of a helicopter trauma service. Also, the effect of the service on the health status of the surviving patients may be important for the cost-effectiveness profile of the service. However, the latter item could not be included in the model, because no data on this effect were available. The modelling approach was chosen to determine whether a clinical trial would be required, because a clinical trial in this field was expected to be time consuming, costly, and difficult to realise because of the difficulty in obtaining a reliable control group. Here, the modelling approach did not lead to a definite conclusion on the cost-effectiveness of the service, but led to a recommendation to focus further research on specific items. It also showed that there was no scientific basis for supporters or adversaries of the helicopter service, who both claimed no further research was necessary because the service

obviously would - respectively wouldn't - be cost-effective, to sustain either claim without a clinical trial. Therefore, eventually a clinical trial was designed and performed.

The construction of a clinical decision rule for the diagnosis of pulmonary embolism (chapter 2) illustrates that modelling techniques can be applied to obtain optimal information from clinical observations and to develop diagnostic tools. The multivariate logistic regression model developed for the diagnosis of pulmonary embolism contains only 5 items - 2 items that can be obtained in the interview of the patient (previous deep venous thrombosis and recently developed or worsened cough), 2 items related to the physical examination (wheezing and body temperature above 37°C) and 1 item containing a qualitative description of the standard diagnostic test for these patients (multiple defects on the perfusion scan).[2] If no modelling technique would have been used, only the clinical assessment of the physician would have been available to assess the probability of pulmonary embolism without further diagnostic testing. The final model showed a better sensitivity and specificity than the clinical assessment by the physician (before lung-scanning). Microsimulation showed that models based on the selected combination of variables tended to be fairly robust. But, even though the sensitivity and specificity of the model seem favourable in the derivation phase presented in chapter 3, if the model is to be used not only for analytic purposes but also as a diagnostic tool in clinical practice it will still have to prove its usefulness in a prospective validation phase. There are two main reasons why the clinical decision rule may not perform as well as expected in this prospective validation phase, and eventually in clinical practice. First, physicians may score the items included in the model differently if they know the items will have a significant contribution to the result of the clinical decision rule. In other words they may - subconsciously - try to direct the outcome of the decision rule towards their own intuitive opinion. Secondly, a change in distribution of the variables and their outcomes in the future patient population may lead to invalidity of the rule. This may especially be the case for rules based on small patient populations.

The chapter concerning the evaluation of the cost-effectiveness of alternative strategies (chapter 3) shows that computer modelling technology can help to choose the most cost-effective strategy from more than 10,000 possible diagnostic strategies. Without the use of modelling techniques only the analysis of a few, predefined, diagnostic strategies would have been possible, whereas now an indication can be given of the costs and effects of the whole range of possible

diagnostic strategies. Data on sensitivity and specificity of several new diagnostic tests for pulmonary embolism or venous thrombosis were obtained from a prospective cohort study in 452 Dutch patients and incorporated in a decision tree model. The decision tree model was built in a spreadsheet programme and Bayes' theorem was used to combine the sensitivity and specificity of the separate diagnostic tests into a combined sensitivity and specificity for the combination of tests.[3,4,5] The main assumption that had to be made in the analysis - for technical reasons - was that of the conditional independency of the test results. In this case, the assumption did not seem to have large effects since the application of the optimal strategy to the original database did not lead to results that were very different from those predicted by the model. However, further research into this field should be done to see whether simple methods can be found to analyse strategies containing tests with dependent results in a similar way. In a subsequent RCT the costs and effects of the diagnostic strategy selected by modelling may be prospectively compared to those of the prevailing strategy.

The use of modelling techniques as *substitute* for (parts of) trials that are thought to be unethical, unfeasible or impossible was illustrated with an example in which the survival of the control group was modelled, because of the absence of a contemporary control group. This example (chapter 4) shows that modelling permits reliable evaluation of medical technologies in this situation. In the cost-effectiveness analysis of liver transplantation the alternative would have been either to conduct a randomised controlled trial, which was thought to be unethical because liver transplantation was already assumed to be effective abroad, or to use a historical control group. The optimal solution for the assessment of the effectiveness of transplantation was found in the use of historical data incorporated in a Cox regression model. In the analysis, Cox regression models were used to predict survival probability for each patient as it would have been without transplantation.[6] However, it should be kept in mind that only models based on patients comparable to the patients analysed and models from countries and times with a level of care similar to the current level of conservative treatment can be used to replace a contemporary control group.

Chapter 5 shows that modelling techniques can be used for the *extrapolation* of the effects of a clinical trial, in this case to another country and another time perspective. The effects of ACE-inhibitors on cardiovascular morbidity of patients with a-symptomatic heart failure after a myocardial infarction, which were assessed in a North American trial,[7] were extrapolated to the Dutch situation and to a follow-

up of 20 years in a Markov model.[4,8,9] Because no data on costs were published for the original trial the trial was also *complemented* with cost data specific of the Dutch situation. Although ACE-inhibitors were used as a preventive therapy, a RCT was feasible because the treatment showed effects within a reasonable time-period. In this example the alternative would have been to repeat the clinical trial as a cost-effectiveness trial in the Netherlands, which would have been costly, time-consuming, and - as the results of the modelling analysis show - probably unnecessary.

In the extrapolation and complementation of the trial results many uncertainties had to be incorporated into the model. The effect of these uncertainties on the final outcome was analysed with a multivariate sensitivity analysis in a microsimulation model,[10,11,12,13] which was superimposed on the Markov model. Here also the assumption of independency of the effects of the separate variables had to be made for technical reasons. The sensitivity analysis led to a frequency distribution of cost-effectiveness ratios, which included the "worst outcome scenario", combining maximum costs with minimal effects, health care planners might be interested in before approving the introduction of such preventive treatment. This shows that multivariate sensitivity analysis, which used to be almost impossible in situations where more than a few variables were deemed to be important, can easily be obtained by modelling techniques, under the assumption of independency of the effects of the separate variables.

The use of modelling techniques to predict the consequences of the *implementation* of a new technology was illustrated with examples on the assessment of macroeconomic consequences of the introduction of ACE-inhibitors (chapter 5) and on the planning of ECMO facilities (chapter 6).

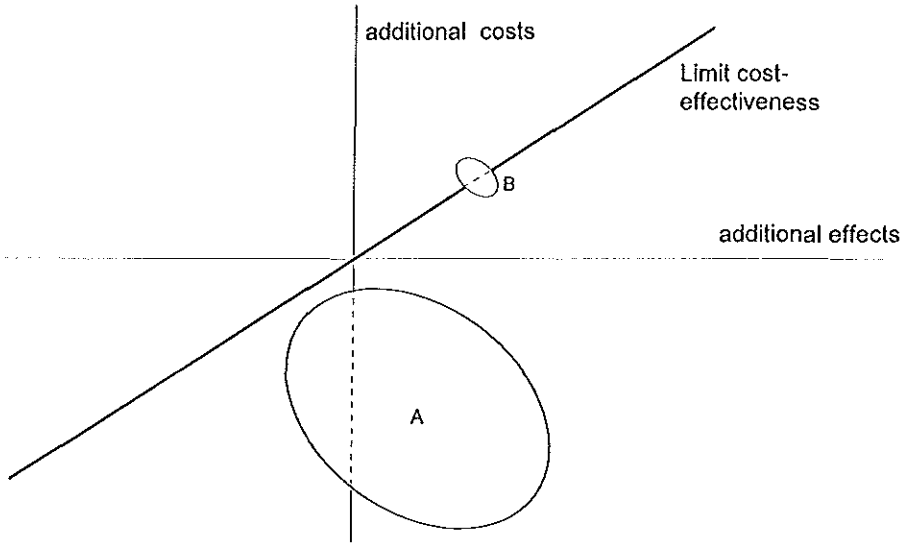
The use of modelling in the analysis of the macro-economic consequences of the introduction of a new technology was described in the chapter on the introduction of preventive treatment with ACE inhibitors (chapter 5). The Markov model indicates the nation-wide annual additional costs of this treatment will probably increase during the first years after the introduction of ACE-inhibitors as preventive therapy,[4,8,9] until, after some 40 years, a steady state might be reached in which the annual additional costs amount to approximately 42 million Dutch Guilders. This last figure should not be taken as the absolute truth but only as a frame of reference, because new developments and therapies during those 40 years are, of course, not included in the model. Despite these uncertainties,

modelling may give an impression of the nation-wide consequences even before the treatment is implemented, thereby saving time and money.

Finally, the example on the planning of neonatal ECMO facilities shows that an efficient planning of health care facilities can be supported by modelling techniques. The model, a microsimulation,[4,13,14] describes the duration of ECMO treatment, which was randomly drawn from the observed frequency distribution of treatment durations in the Netherlands, for different numbers of patients requiring treatment annually and for different numbers of treatment centres. With the model the numbers of patients that will have to be referred to centres outside the region are assessed. The model showed that in the Netherlands no additional ECMO facilities will be needed in the near future, if between 2 and 3 referrals to centres abroad are thought to be acceptable and feasible. This example illustrates the potential role of modelling techniques in the planning of health care facilities. Whereas such techniques are commonly used in operations research for industrial purposes, their use in health care has hitherto been limited. However, with a more formalised planning uneconomic proliferation of health care facilities may be prevented and a more efficient level and distribution of care may be obtained.

In this thesis the place of mathematical-statistical models in medical technology assessment is discussed and illustrated with examples. A large advantage of modelling techniques over RCTs not discussed in detail in the previous chapters, is that they allow for a better incorporation of rare events than RCTs do. For example, in the evaluation of a relatively inexpensive new therapy for oesophagus varices one liver transplantation occurring (by chance) in one of the arms of the RCT may have a large influence on the costs of that arm, whereas by modelling the occurrence of liver transplantation may be brought to realistic proportions in both treatment groups.

Figure 1 *The 95% confidence intervals for two medical technologies A and B. Technology A can be called cost-effective despite the large confidence interval, whereas no conclusions can be drawn about the cost-effectiveness of technology B.*



The overview thus far gave an optimistic impression of the many and varied opportunities modelling techniques may offer in this field. However, a warning should be given: models tend to start a life of their own and may suggest a reliability that does not exist: models can never be more reliable than the data they are based on. In addition to this there are other limitations to modelling techniques:

- Unfortunately, in general, modelling is viewed with more suspicion than RCTs, because of the incorporation of data from different sources which introduces additional bias and uncertainty. However, figure 1 illustrates that a large uncertainty margin need not represent a problem for decision making as long as every point within the uncertainty area surrounding the cost-effectiveness of the technology is on one side of the line representing the societally acceptable level of cost-effectiveness (situation A).[15] In contrast, a much smaller 95% confidence interval which passes the limit of cost-effectiveness (situation B) can definitely prohibit the drawing of conclusions from assessment B. Therefore, as long as the limit of cost-effectiveness is not passed - and as long as the decision based

on the analysis remains unaltered - even modelling techniques leading to large uncertainty margins may still return useful results.

- Modelling techniques can only be used if enough representative data are available. Even though in theory an unlimited number of wild guesses can be made and incorporated into a model, in practice a majority of the assumptions incorporated into a model need some basis in reality to give the model validity.
- Technical limitations mainly lay in the assumption about conditional independency of the results of tests or of the influence of variables on the outcome. Computer time, which used to be the main limitation to the application of modelling techniques, is hardly a limiting factor any more with modern computer facilities.

In conclusion, it can be stated that mathematical-statistical modelling certainly deserves a place within the distinct phases of medical technology assessment. This thesis shows that modelling can lead to useful results in, amongst others, the design of trials, the assessment of the effectiveness of medical technology in situations where a contemporary control group is lacking, the improvement of the use of clinical data, the evaluation of the cost-effectiveness of diagnostic strategies, the complementation of trial results with cost data, the extrapolation of trial results to other countries and time horizons, the evaluation of macro-economic consequences of the introduction of medical technology and in the planning of health care facilities. In medical technology assessment modelling should not be seen as a strict replacement for randomised controlled trials. In some situations a classic RCT still remains the optimal analysing technique. But in others, modelling may return information much faster and cheaper than a randomised controlled trial ever could.

In general it should be kept in mind that the usefulness of models is limited through the assumptions and data included. Therefore, as a recent editorial in the New England Journal of Medicine on the use of models in cost-effectiveness studies stipulated, in the reporting of the results of a modelling analysis all data, all assumptions and any model used should be included and explained clearly.[16] Because of the large uncertainty margins often surrounding the outcome of models, the requirement to perform a detailed, preferably multivariate, sensitivity analysis should be added to this.

References

1 Rivett P. The craft of decision modelling. Chichester: John Wiley & Sons, 1994

2 Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons, 1989

3 Sox HC, Blatt MA, Higgins MC, Marton KI. Medical Decision Making. Boston: Butterworth-Heineman, 1988

4 Keeler E. Decision trees and Markov models in cost-effectiveness research. In: Sloan FA (editor). Valuing health care. Costs, benefits, and effectiveness of pharmaceuticals and other medical technologies. Cambridge: Cambridge University Press, 1995; 185-206

5 Weinstein MC, Fineberg HV, Elstein AS, Frazier HS, Neuhauser D, Neutra RR, McNeil BJ. Clinical Decision Analysis. Philadelphia: WB Saunders, 1980

6 Hop WCJ, Hermans J. Statistische analyse overlevingsduren. T soc Geneesk 1981; 59(8): 279-288

7 Pfeffer MA, Braunwald E, Moyé LA, et al. Effect of captopril on mortality and morbidity in patients with left ventricular dysfunction after myocardial infarction. Results of the Survival and Ventricular Enlargement trial. New England Journal of Medicine 1992; 327: 669-677

8 Sonnenberg FA, Beck JR. Markov models in medical decision making. A practical guide. Medical Decision Making 1993; 13: 322-338

9 Beck JR, Pauker SG. The Markov process in medical prognosis. Medical Decision Making 1983; 3(4): 419-458

10 Doubilet P, Begg CB, Weinstein MC, Braun P, McNeil BJ. Probabilistic sensitivity analysis using Monte Carlo Simulation. A practical approach. Medical Decision Making 1985; 5(2): 157-77

11 Critchfield GC, Willard KE. Probabilistic analysis of decision trees using Monte Carlo simulation. Medical Decision Making 1986; 6(2): 85-92

12 Hazen GB. Stochastic trees. A new technique for temporal medical decision modelling. Medical Decision Making 1992; 12(3): 163-177

13 Klein RW, Dittus RS, Roberts SD, Wilson JR. Simulation modeling and health-care decision making. Medical Decision Making 1993; 13(4): 347-354

14 Hout BA van, Oortmassen GJ van. Het modelleren van chronische ziekten. Enkele alternatieven. T Soc Gezondheidsz 1992; 70: 362-67

15 Hout BA van, Al MJ, Gordon GS, Rutten FFH. Costs, effects and C/E-ratios alongside a clinical trial. Health Economics 1994; 3: 309-319

16 Kassirer JP, Angell M. The Journal's policy on cost-effectiveness analyses [editorial]. New England Journal of Medicine 1994; 331(10): 669-670

SAMENVATTING

In dit proefschrift wordt de plaats van wiskundig-statistische modellen binnen de Medische Technology Assessment (MTA) geïnventariseerd en besproken. MTA is een vorm van onderzoek waarbij de verschillende aspecten van een medische technologie in kaart worden gebracht om de besluitvorming omtrent die technologie zo goed mogelijk te laten plaatsvinden. Kosten-effectiviteitsanalyse - waarin de kosten die aan de technologie verbonden zijn worden gerelateerd aan de effecten op overleving, morbiditeit en kwaliteit van leven - vormt meestal de kern van een dergelijk onderzoek. Daarnaast kunnen echter ook andere aspecten, bijvoorbeeld op het juridische of ethische vlak, in de analyse worden betrokken.

Vaak wordt gerandomiseerd (dubbel blind) onderzoek, waarbij de te onderzoeken technologie wordt vergeleken met de beste op dat moment gangbare behandeling, aanbevolen als de meest geschikte methode om de kosten en effecten van een medische technologie aan elkaar te relateren. Dergelijk onderzoek is echter niet altijd mogelijk of wenselijk. Indien er (nog) geen gerandomiseerd onderzoek kan plaatsvinden kunnen modelleringstechnieken worden gebruikt om toch een indruk van de kosten en effecten van de technologie te krijgen. Deze technieken kunnen in alle fasen van MTA worden toegepast: bij de voorbereiding en ontwikkeling van een kosten-effectiviteitsstudie, als vervanging van (gedeelten van) het gerandomiseerd onderzoek, om kostengegevens aan een gerandomiseerd onderzoek toe te voegen, bij de extrapolatie van onderzoeksresultaten naar een langere termijn, een ander land of een andere patiëntenpopulatie en bij de implementatie van onderzoeksresultaten. De verschillende toepassingen zullen hieronder kort worden besproken.

Bij de *voorbereiding en ontwikkeling* van een kosten-effectiviteitsstudie kan modellering worden gebruikt om te onderzoeken of gerandomiseerd onderzoek nodig en zinvol is. Bovendien kan een indruk worden gekregen van de onderwerpen die in ieder geval in een dergelijk onderzoek dienen te worden betrokken. Deze toepassing wordt geïllustreerd met een voorbeeld waarbij de kosten en effecten van de introductie van helikopters met traumateams in

Nederland worden onderzocht (hoofdstuk 1). In de analyse werden alle op het moment van het onderzoek voor Nederland beschikbare relevante gegevens gecombineerd in een model. Uit deze benadering bleek dat de onzekerheid rondom de overlevingswinst als gevolg van de helikopterservice en de kosten van een ziekenhuisopname voor een ernstig gewonde patiënt van veel groter belang waren voor de uitkomst dan de onzekerheid rondom de kosten van de helikopterservice zelf. Ook bleek dat er op basis van de beschikbare gegevens geen duidelijke uitspraken te doen waren over de kosten-effectiviteit van de service. Uiteindelijk werd derhalve besloten een experimenteel onderzoek uit te voeren om tot een betrouwbare uitspraak omtrent de kosten-effectiviteit van de helikopterservice in Nederland te komen. Door de modelmatige voorstudie was duidelijk dat analyse van de effecten op de overleving en de kosten van ziekenhuisopname een belangrijke plaats in het experimentele onderzoek verdienden.

Ook kan modellering gebruikt worden om methoden te ontwikkelen die bij het gerandomiseerd onderzoek gebruikt kunnen worden. Deze toepassing wordt geïllustreerd met een voorbeeld waarin de klinische informatie die nodig is om de diagnose longembolie te stellen met behulp van logistische regressie wordt gecombineerd tot een klinische beslisregel (hoofdstuk 2). De relatief eenvoudige beslisregel suggereerde een betere sensitiviteit en specificiteit dan de klinische indruk van de arts en kan, na een prospectieve validatie, mogelijk zelfs in de praktijk worden gebruikt als een diagnostische test. In een gerandomiseerd onderzoek naar de beste strategie bij de diagnostiek van longembolieën kan de beslisregel worden gebruikt om klinische gegevens in het onderzoek te betrekken.
Bij het selecteren van de strategieën die in het gerandomiseerde onderzoek worden opgenomen maakt modellering, in combinatie met een cohort-onderzoek, een snelle evaluatie van de kosten en effecten van een groot aantal strategieën mogelijk. Dit wordt geïllustreerd met de evaluatie van de kosten en effecten van diagnostische strategieën bij patiënten met een mogelijke longembolie (hoofdstuk 3). In dit voorbeeld worden, in een combinatie van een besliskundige analyse en een prospectief cohort-onderzoek bij 452 patiënten, de kosten en effecten van alle mogelijke diagnostische strategieën (meer dan 10.000!) met elkaar vergeleken. De evaluatie van een dergelijk groot aantal strategieën zou met behulp van gerandomiseerde trials natuurlijk onmogelijk zijn geweest. Uiteindelijk kan de meest kosten-effectieve strategie, geselecteerd met deze methode, vergeleken worden met de gangbare strategie in een gerandomiseerd onderzoek.

Modellering kan ook gebruikt worden als *vervanging* van een (gedeelte van) gerandomiseerd onderzoek.
Bij de kosten-effectiviteitsanalyse van screeningsprogramma's en preventieve behandelingen kan modellering worden gebruikt om gebeurtenissen in de verre toekomst of zeldzame - maar belangrijke of kostbare - complicaties in de analyse te betrekken. Gerandomiseerd onderzoek zou in dit geval meestal kostbaar en langdurig zijn.
Bij de evaluatie van behandelingen die levensreddend kunnen zijn bij patiënten waarvoor eigenlijk geen alternatieve behandelingsmogelijkheden bestaan, wordt gerandomiseerd onderzoek snel onethisch gevonden. Door modellering van de controlegroep kan in dergelijke gevallen toch een uitspraak gedaan worden over de (kosten-)effectiviteit van de behandeling. Deze toepassing wordt geïllustreerd met een onderzoek naar de kosten en effecten van levertransplantatie, waarbij de overleving van de controlegroep werd gemodelleerd met behulp van Cox regressie (hoofdstuk 4). Op deze wijze kon aannemelijk worden gemaakt dat levertransplantatie een aanzienlijke overlevingswinst kan bewerkstelligen ten opzichte van conservatieve behandeling.
Naast bovengenoemde indicatie zijn ook andere toepassingen van modellering ter vervanging van (gedeelten van) gerandomiseerd onderzoek denkbaar. Zo kan modellering worden toegepast om een kosten-effectiviteitsanalyse te doen van technologieën die gebruikt worden bij relatief zeldzame ziekten of ziekten met een lage incidentie, zoals bijvoorbeeld bij de kosten-effectiviteitsanalyse van inenting na hondsdolheidbesmetting.

Indien een gerandomiseerd onderzoek beschikbaar is, waarbij geen kosten werden verzameld, kan modellering worden gebruikt deze alsnog aan het onderzoek *toe te voegen*, zodat alsnog een kosten-effectiviteitsratio kan worden berekend. In dit proefschrift is deze methode toegepast bij een kosten-effectiviteitsanalyse op basis van de gepubliceerde gegevens van het SAVE onderzoek, een groot Amerikaans onderzoek naar de effecten van ACE-remmers als preventieve behandeling na een hartinfarct (hoofdstuk 5).

Modellering kan ook worden gebruikt om een meer betrouwbare kosten-effectiviteitsanalyse mogelijk te maken door de *extrapolatie* van de resultaten van gerandomiseerd onderzoek. In een model kunnen de resultaten van vrijwel identieke trials worden gecombineerd of kunnen de intermediaire uitkomstmaten van een trial (zoals het aantal maagzweren dat wordt voorkomen met een bepaalde medicatie) worden vertaald naar uitkomstmaten die bruikbaar zijn voor

kosten-effectiviteitsanalyse, zoals gewonnen levensjaren en voor kwaliteit van leven gecorrigeerde gewonnen levensjaren.

In dit proefschrift wordt het gebruik van modellering bij de extrapolatie van onderzoeksgegevens geïllustreerd met Markov model waarmee de resultaten van het SAVE onderzoek (hoofdstuk 5) zijn geëxtrapoleerd naar Nederland en naar de lange termijn (van 4 jaar naar 20 jaar). Omdat in de extrapolatie werd vergeleken met de gangbare Nederlandse praktijk in plaats van met de oorspronkelijk placebobehandeling is er ook sprake van een extrapolatie van een door de inclusiecriteria van het onderzoek beperkte werkzaamheidsanalyse naar een effectiviteitsanalyse.

Tenslotte kunnen modelleringstechnieken ook worden gebruikt bij beslissingen rond de *implementatie* van medische technologieën.
Onderzoeken naar de gevolgen van nationale implementatie zijn in de praktijk vaak kostbaar en langdurig. Bovendien is het vrijwel onmogelijk om de implementatie na een nationale introductie alsnog te stoppen als de resultaten ongunstig uitvallen. Met behulp van een Markov model werden de macro-economische gevolgen gedurende de eerste 50 jaar na introductie van preventieve behandeling met ACE-remmers in Nederland geschat (hoofdstuk 5). Hoewel deze berekening natuurlijk met onzekerheid omgeven is, kan deze methode een indruk geven van de te verwachten effecten voordat daadwerkelijk introductie op nationale schaal plaatsvindt.
Modelleringstechnieken kunnen ook worden gebruikt bij de planning van nieuwe medische faciliteiten. Het laatste voorbeeld toont aan dat met behulp van een microsimulatie voor ECMO-behandeling bij pasgeborenen kan worden geschat hoeveel kinderen er jaarlijks voor behandeling naar het buitenland moeten worden verwezen bij een bepaald aantal behandelingsfaciliteiten in Nederland (hoofdstuk 6). Het resultaat van deze analyse is dat er in Nederland geen extra ECMO-faciliteiten voor pasgeborenen nodig zijn als 2 tot 3 verwijzingen naar het buitenland per jaar acceptabel wordt gevonden. Modellering leidt hierbij sneller en goedkoper tot resultaten dan een evaluatie die pas plaatsvindt nadat reeds een groot aantal faciliteiten is opgericht.

Bovenstaande voorbeelden tonen aan dat modelleringstechnieken, zoals Markov processen, microsimulatie en beslisbomen, een ruime toepasbaarheid hebben binnen MTA-onderzoek. Daarnaast kennen ze natuurlijk ook beperkingen. Deze zijn voornamelijk gelegen in de betrouwbaarheid, beschikbaarheid en

relevantie van de gegevens die in het model zijn opgenomen, de grote onzekerheidsmarge die modellering soms met zich mee brengt en in technische opzicht in de gecompliceerdheid van de modellering bij conditioneel afhankelijke invloed van variabelen op de uitkomst, en - in steeds mindere mate - in de beschikbaarheid van computerfaciliteiten.

De conclusie van het hier gepresenteerde onderzoek is dat er zeker een plaats is voor modellering in de verschillende fasen van Medische Technology Assessment, mits de onzekerheden en de veronderstellingen die aan het gebruikte model gerelateerd zijn duidelijk gedocumenteerd worden in de verslaglegging van de resultaten. Een uitgebreide, liefst multivariate, gevoeligheidsanalyse is daarbij gewenst.

Curriculum Vitae

Bowine Michel was born in Leiden on January 22nd, 1962. After her graduation from the Gymnasium-β section of the Gymnasium Erasmianum in Rotterdam in 1979, she went to Medical School in Leiden. As a student she worked at the department of Anatomy, where she did a research project on methods to measure changes in faces of small children with a cleft lip and / or palate, and at the department of Infectious Diseases, where she researched the infuence of γ-interferon on the bacterial killing activity of peritoneal macrophages in mice. She also did three months of practical work at the department of Immunohematology on the role of HLA-DR positive T-lymphocytes in the mixed lymphocyte culture. In January 1987 she got her Medical Degree, cum laude.

After her studies she first did a meta-analysis of literature on antibiotic treatment in Cystic Fibrosis patients for an international workgroup on this subject. In May 1987, she started her training in Internal Medicine in the Red Cross Hospital in The Hague. After she had decided not to finish this training by the end of 1988, she worked as a member of staff at the laboratory for Virology of the Municipal Health Service in Rotterdam.

In 1990 she became a Research Fellow at the institute for Medical Technology Assessment of the Erasmus University Rotterdam, detached at the institute of Public Health. Since July 1994 she works for the institute for Medical Technology Assessment at the institute of Health Policy and Management of the Erasmus University.

In 1993 she was registered as epidemiologist A and in September 1994 she got her propaedeutics in Business Economics at the Open University in Heerlen, the Netherlands.

Acknowledgements

There are several people I'd like to mention:

Ben van Hout, my co-promotor, without whom this thesis would not exist. He was involved in almost every aspect, from the very beginning to the very end. Working with him is a pleasure because he often comes up with creative solutions for seemingly unsolvable problems. If only that phone wouldn't ring so often.....
Frans Rutten, my promotor, who was involved in the later stages of this thesis. His useful and detailed comments certainly improved it.
The co-authors of the papers included in this thesis, especially Harry Büller with whom I had many pleasant discussions over the years.
The colleagues in my cluster at the institute for Medical Technology Assessment: Ben, Maiwenn, Rob, Joseph, Jan and Renate.
The other colleagues of the institute for Medical Technology Assessment, the institute of Health Policy and Management and the institute of Public Health, including, of course, the secretaries, administrators and computer-assistants .
The ECMO and MALEDIVE research teams at the department of surgery of the Sophia Kinderziekenhuis Rotterdam, the department of neonatology of Nijmegen University Hospital and the department of Hemostasis, Thrombosis, Atherosclerosis and Inflammation Research of the Academic Medical Centre.
Dr C Vermeij-Keers who first gave me the idea to write a PhD thesis. She offered me a job to write one at the department of Anatomy in Leiden already some 15 years ago. At that time I thought there were more important things than a PhD thesis. Now, of course, I know better.
Loes, Jan-Gerben, Caroline and Mary-Lou with whom I shared my room in the office over the years. They helped me through some arduous moments.
Hans en Saskia, my "paranymfen" and best friends.
My mother who thoroughly cleared the final version of this thesis from all inconsistencies.
My father who was always there with helpful advise along the way.
And all friends who morally supported me in these, sometimes difficult, years.

Thank you all.

"Would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to," said the Cat.

"I don't much care where ....." said Alice.

"Then it doesn't matter which way you go," said the Cat.

".....so long as I get *somewhere*," Alice added as an explanation.

"Oh, you're sure to do that," said the Cat, "if you only walk long enough."

Lewis Carroll: Alice in Wonderland